PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 18 February 15, 1932 Number 2

# *THE DEVELOPMENTAL MORPHOLOGY OF INFANT BEHAVIOR PATTERN*

*(With a Demonstration of Methods of Systematic Cinematography)*

By Arnold Gesell, M. D.

Director of the Clinic of Child Development, School of Medicine, Yale University

Read before the Academy November 17, 1931

The behavior of the human infant is fundamentally ordered by ascertainable, though poorly understood, laws of growth. Behavior is molded into characteristic forms by virtue of these laws of growth. And such forms, whether statically or dynamically regarded, may be investigated from the standpoint of a developmental morphology.

From the first movements of the foetus to the assumption of erect posture the action systems of the organism display patterning. There are patterns of specific reflex behavior, patterns of totalized, integrated response and patterns of ontogenetic sequence. There is no such thing as utter randomness in infant behavior. Even the sprawling of the eight-weeks infant is distinctive from that of the twelve-weeks-old. This distinctiveness arises out of a deep-seated developmental difference in the patterns of posture and of perception. The organization of the growing complex of infant behavior is faithfully revealed in outward forms of posture, of perceptual adjustment, prehension and adaptive manipulation.

It is possible to capture these visible behavior forms by means of photography. Fortunately we have in the motion picture camera an effective instrument which serves at once for systematic recording and for the experimental analysis of developing behavior. The camera records completely and impartially; and it remembers infallibly. In any given behavior situation it preserves in totality the correlated plexus of the infant's movements. It registers the simultaneous attitude of head, trunk, arms, legs, eyes, fingers, face. It portrays each successive moment of behavior in all its significant togetherness, and permits us to study any moment as a frozen section of behavior pattern. By close seriation of moments, the cinema registers the sequence within a behavior episode. It also

chronicles the developmental changes in pattern at successive age levels. It provides an authentic record which may be subjected to objective and comparative analysis.

Each record incorporates its own time and space values which can be quantitatively expressed. Four feet of 16 mm. film embody 160 individual frames which depict 160 phases of a behavior episode, 10 seconds in duration. Reinstating systematically a similar episode at six advancing age levels yields a developmental record of more than a thousand pattern phases. The photographic data do fair justice to the wealth and complexity of the problems. Cinematography thus becomes a systematic tool for recording the phases of behavior pattern in their immediate dynamic sequence; and for delineating with advancing age the progressive differentiation of behavior patterns in a developmental cycle.

With the aid of the accompanying photographs it will be possible to briefly indicate the methods which have been developed at the Yale Clinic of Child Development, for the cinematographic inventory and charting of the growth of behavior patterns in the human infant. Our subjects are normal infants selected as to race, gestation period, health, parentage and socio-economic factors. Twenty-five or more children have been examined at lunar month intervals from 4 weeks through 56 weeks of age. The infants were observed in a specially designed clinical crib, under controlled conditions. The data include stenographic and cinema records covering the fields of posture, locomotion, prehension, manipulation, attentional regard, exploitive and adaptive behavior. This behavior survey has been in progress for five years and has been accomplished through coördinated studies by research members of the staff of the Clinic.*

The basic arrangement for securing the cinema records is an observation dome 12 feet in diameter located in the photographic laboratory. (Fig. 1) The structure is hemispherical, but is the reverse of its astronomical prototype because the route of observation is from without inward. The operator, observers and parent are stationed outside; the infant is in the crib in the universal focal area of the cameras located on two quadrants which are at right angles to each other. The dome is encased in a white-painted, 16-mesh wire screen and is illuminated with well-diffused soft light by Cooper Hewitt combined with Mazda lamps. The surrounding laboratory is darkened. Through these optical conditions the wall of the dome becomes a one-way vision screen, effectively concealing the observers from the infant who himself is clearly perceived as he displays his repertoire of behavior in his crib.

The clinical crib (Fig. 2) is designed to give uniformity to the experimental behavior situations. It is equipped with adjustable platform, panels and side rails, and a container for behavior test materials.

The infant may be put in a small chair (morris type) before a table top on which a stimulus object is placed in accordance with standardized procedure.

Cinema cameras, spring-wound and housed in sound-quieted boxes, travel on the quadrants and may be poised in any desired position between horizon and zenith. These cameras operate singly; or simultaneously, to give pictorial control of the third dimension. A pair of still cameras in electrical connection may be used to give a simultaneous record. Figure 3 pictures a synchronous frontal and overhead view of a pair of identical twins, each of whom is making approach on a small pellet (7 mm. in diameter) placed on the table top. Incidentally it may be noted that these infant twins were not only extremely similar in physical characteristics, but showed an amazing degree of resemblance in the configuration and the maturing of their behavior patterns.

The cinema film supplies at once the vehicle and the basic material for the analysis of a restricted behavior pattern like the prehension of a pellet; or of a more complicated pattern like the adaptive placement of a pellet in a bottle. A strip of 16 mm. film, two and a half feet long, reconstitutes six seconds of behavior in 100 successive frames. The pattern may be minutely analyzed by inspection of each individual frame. This is accomplished by a movieola, or, more conveniently, by a special projection desk [designed by Halverson] operated by hand, permitting the image to be repeatedly viewed at will in full and slowed motion, or in stilled pattern phases.

The systematic cinema films of The Clinic of Child Development have been classified and catalogued by library methods and are booked in 100-foot reels. Fourteen age levels are represented in the material which because of its flexibility permits a wide range of manipulation for purposes of comparative observation. In developmental studies it is highly useful to bring past and present data into close conjunction. Indeed, the scientific value of cinematography lies in the fact that it renders possible three kinds of observation: analytic, repetitive and manipulative. It also permits to an unlimited extent a genetic codification of the changing patterns of infant behavior. The problem of systematic cinematography is to select and to seriate pattern phases in a manner which will delineate significantly the increments and configurations of behavior growth.

The manner in which the growth changes of behavior pattern assert themselves is illustrated by the group of photographs which show the typical prehension reactions of an infant in the pellet situation at six stages of development (Fig. 4). The characteristic course of the development of pellet prehension patterns is recapitulated below. The following normative synopsis covers 12 lunar-month periods beginning at 8 weeks:

8 weeks: No visual regard for pellet
12 weeks: Transient regard for pellet (rarely)
16 weeks: More prolonged regard, usually delayed
20 weeks: Immediate, definite regard, sometimes with increased hand-arm activity
24 weeks: Approaches pellet with pronate hand; contacts pellet with little or no finger adjustment
28 weeks: Approaches pellet with raking flexion of fingers, without thumb opposition; occasional delayed palmar prehension
32 weeks: Approaches pellet with raking flexion but with increased thumb participation and digital prehension
36 weeks: Approaches and contacts with simultaneous flexion of fingers: prehends with defined thumb and index opposition
40 weeks: Approaches with all fingers extended; contacts with index finger and later prehends by drawing index finger against thumb
44 weeks: Promptly prehends with index and thumb and with increased obliquity of hand attitude
48 weeks: Approaches with index finger extended and lateral digits flexed; prehends with delimited plucking by index and thumb
52 weeks: Approaches and plucks pincerwise with increased deftness

The above summary, though sketchy and condensed, is presented to show in principle the possibility of specifying developmental increments of behavior pattern at lunar month intervals in the first year of life. Behavior patterns when carefully defined become biological indices of the developmental maturity of the infant. With the aid of cinemanalysis the patterns can be portrayed and arrayed in genetic series to make them serviceable for the diagnosis and interpretation of developmental status.

To demonstrate the significance of a comparative approach in the study of developmental increments, we have prepared a special film picturing an infant's behavior at 24 weeks and again at 28 weeks, and then again both of these "behaviors" of the selfsame infant in simultaneous juxtaposition. This film, illustrating the method of *coincident projection*, was shown in connection with the present paper. Figure 5 represents an excerpt of several frames taken from the duplex section of the film. Each frame consists of two prints made from two originally independent negatives. In actual projection the left-hand print renders the cube behavior of a 24 weeks infant; the right-hand print renders comparable behavior of the selfsame infant at 28 weeks. The rendering on the screen is simultaneous so that we secure an immediate comparative view of two adjacent sets of pattern, expressing two successive levels of maturity. Such a comparative film makes manifest the objective, analyzable and measurable aspects of behavior patterning; and indicates the possibility of building up a morphological account of that "progressive differentiation coördinated in time and place" which is the essence of mental as well as physical growth.

The nature of the developmental increment which occurs between 24

## ILLUSTRATIONS

Figure 1. The Photographic Recording and Observation Dome. The dome is located in the laboratory of The Yale Clinic of Child Development. The interior of the dome is illuminated. The surrounding laboratory is darkened. The infant is examined in the clinical crib in the center of the dome. The dome is encased in a one-way vision screen which permits complete visibility for the observers outside but conceals the observers from the infant. A motion picture camera enclosed in a sound-quieted box attached to the curved track of the dome makes a cinema record of the infant's behavior.

Figure 2. Clinical Crib.

Figure 3. Simultaneous Instantaneous Photographs of a Pair of Identical Twins, showing similarity of behavior pattern in prehensory approach upon a pellet.

Figure 4. Patterns of Pellet Prehension at 20, 24, 28, 32, 40, 52 Weeks.

Figure 5. Frames from a Cinema Film, Illustrating the *Method of Coincident Projection*. The patterns of cube behavior in the same infant at 24 and at 28 weeks are shown in juxtaposition.

Figure 6. Stilled Pattern Phases of the Cube Behavior of an Infant at 24 Weeks.

Figure 7. Stilled Pattern Phases of the Cube Behavior of an Infant at 28 Weeks.

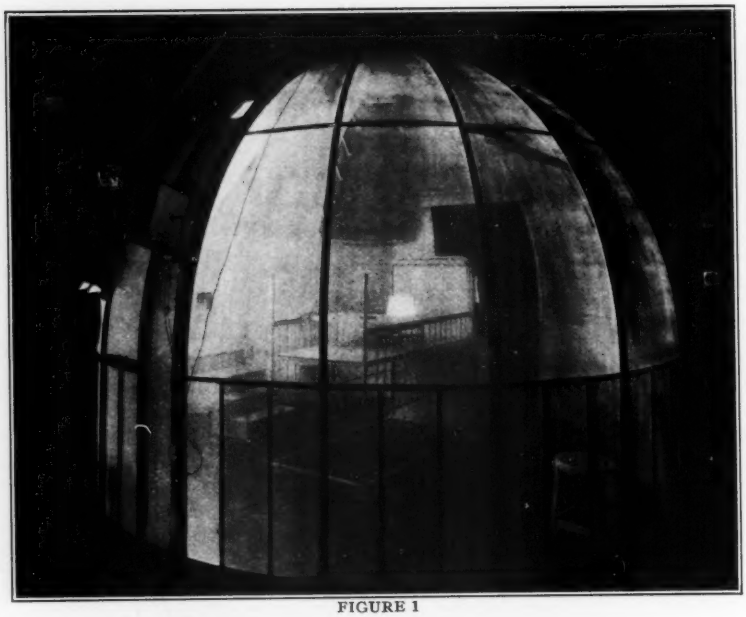


FIGURE 1

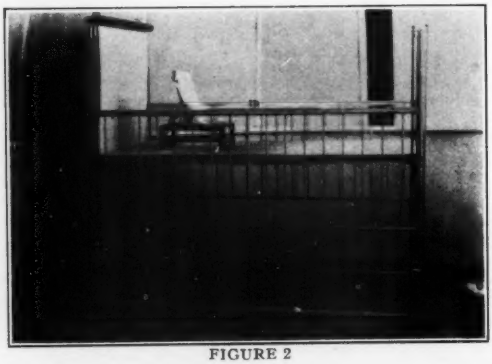


FIGURE 2

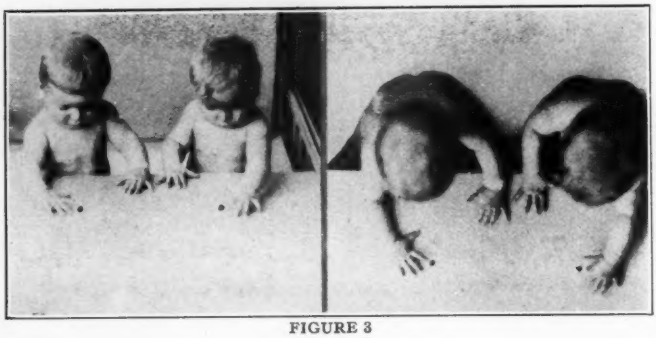


FIGURE 3

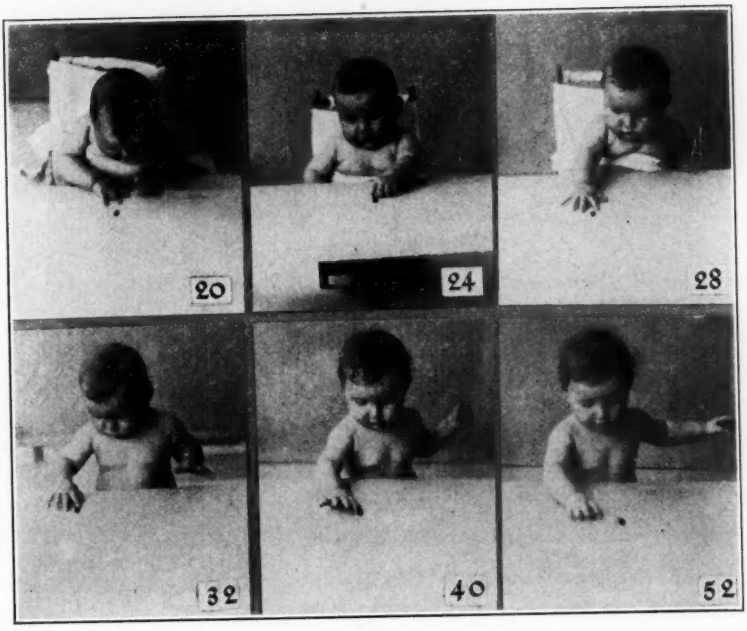



FIGURE 4

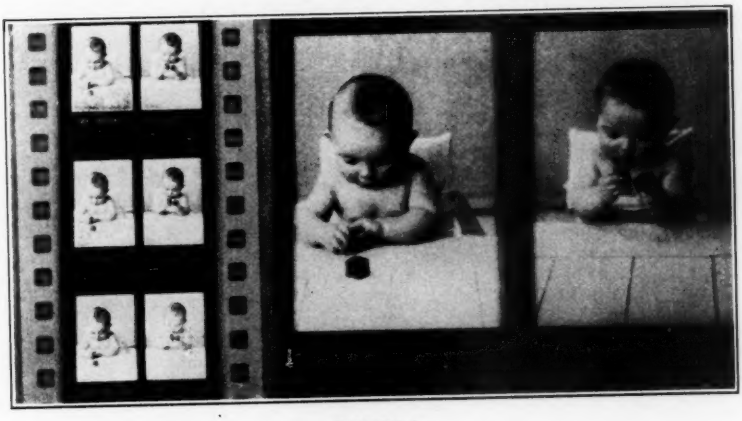
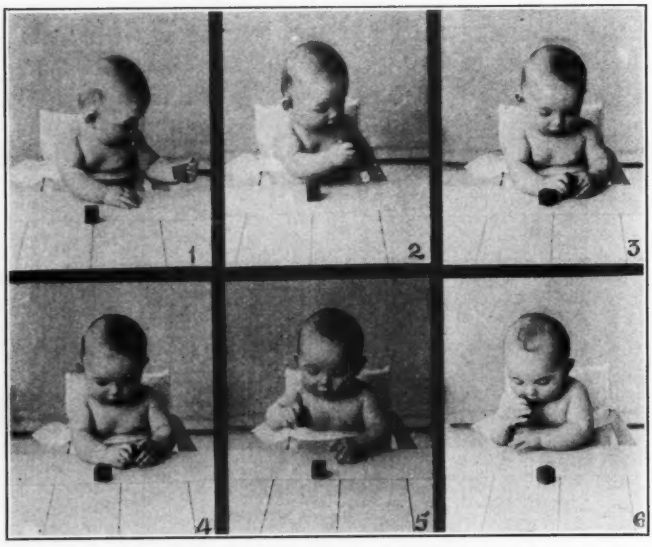
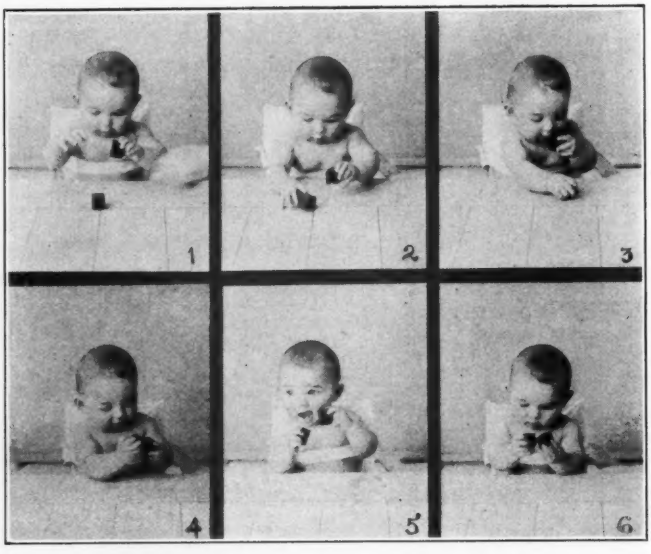


FIGURE 5




24 WEEKS - 4 - 14 - 17.5 - 18 - 20.5 - 36 Secs.

FIGURE 6




28 WEEKS - 10 - 11 - 19.5 - 22 - 26 - 29.5 Secs.

FIGURE 7

and 28 weeks of age is further illustrated in the photographic frames assembled in figures 6 and 7. Each individual photograph represents a stilled pattern phase of a behavior sequence. The behavior episode is initiated by the examiner who places a one-inch red cube in the left hand of the child and then places a similar cube on the table top of the clinical crib within prehensory reach. The time of occurrence of each phase of behavior is indicated in seconds in the legends for both ages. These stilled pattern phases disclose the following differences between the two age levels of this infant: At 24 weeks he chiefly heeds the cube in hand, intently regards this cube, transfers it, bangs it and mouths it. He gives transient regard to the cube on the table. At 28 weeks he first heeds the cube in hand but soon perceives the table cube and promptly prehends it; he retains both cubes, regards each in turn and brings them into sketchy contact with each other. These differences indicate a palpable advance in behavior organization. One of the systematic tasks of developmental psychology is the formulation of such pattern differences by normative and, as far as possible, by biometric methods.

It is not, of course, suggested that all infants pursue precisely the same course of development even in such functions as posture and prehension; but our normative studies have shown significant consistencies in the tempo, the forms and sequences of behavior growth which point to the presence of a lawful developmental mechanics. Systematic cinematography helps us to delineate by ordered specification many of the normative characters and trends of this behavior growth. We have, therefore, begun the compilation of a cumulative Photographic Atlas of Infant Behavior Patterns, embodying objective, normative criteria for the further study of the constant and variable features of mental growth.

Without implying any dualism it is suggested that mental growth like physical growth is a process of morphological organization. The infant does not grow psychologically by a compounding of discrete reflexes, nor solely by a conditioning of innate functions. He is a highly unified organism by virtue of the intrinsic architecture of his nervous system. He is sensitive to environment, but he achieves his behavior capacities primarily through inherent growth processes. His whole behavior equipment, including acquired dynamic traits, is regulated and consolidated by a stable substrate of maturation which keeps casual influences well in check. His psychological make-up, therefore, has a consistent organic structure constituted of action patterns which are basically the product of a lawful morphogenesis.

* The following have participated in the program of research: Dr. Helen Thompson, Research Associate in Biometry; Dr. Henry M. Halverson, Research Associate in Experimental Psychology; and, at the younger age levels, Dr. Catherine Strunk-Amatruda, Research Assistant in Pediatrics.

# ALLOTROPY OF LIQUID BENZENE

By Alan W. C. Menzies and D. A. Lacoss

Frick Chemical Laboratory, Princeton University

Communicated December 19, 1931

Because we have been studying the effect of so-called intensive desiccation upon the behavior of liquid benzene, we were interested in the literature bearing upon the physical properties of this liquid. Ordinary benzene is usually considered to approach the normal rather than the abnormal type of liquid. We wish here to point out that the rate of change with temperature of several physical properties of liquid benzene suffers a more or less abrupt alteration in the neighborhood of 40°. The change is not conspicuously large and, doubtless for this reason, seems to have escaped comment, if not notice. Furthermore, inaccurate measurement may entirely mask the anomaly.

A. On plotting against temperature Sidney Young's[1] experimental values for the specific volume of liquid benzene, we observed that the graph showed a faster change in direction near 36° than at neighboring temperatures. This is shown in figure 1, graph A, where we have drawn two intersecting straight lines through the observed points. These intersect about 36°. If the specific volumes as observed by Meyer and Mylius[2] are similarly graphed, the point of intersection is found several degrees higher.

B. The same figure, graph B, shows in like manner a plot of the values of Parker and Thompson[3] for the refractive index of benzene against temperature. Our straight lines intersect about 43°.

C. On the same figure, graph C exhibits the influence of pressure upon the melting point of benzene, according to the experimental observations of Puschin and Grebenschtschikow.[4] The straight lines drawn by us intersect about 51°.

D. On the same figure, graph D shows values for the dielectric constant, taken directly from the thesis of W. N. Stoops.[5] The values given in the publication of Smyth and Stoops[6] had been interpolated so as to yield rounded temperatures, and do not show the break so sharply. Our straight lines intersect about 42°.

E. On the same figure, curve E shows the values of specific heat as measured by Richards and Wallace[7] in this laboratory. They have kindly permitted us to use their data, which will be published shortly. A similar inflexion is obtained on graphing the (more numerous) observations of Tréhin,[8] near 44°; but not if those of Williams and Daniels[9] are used.

In graphing the above five physical properties we have arranged the direction of the scales so as to avoid crossing of the graphs, since these are

best inspected by viewing them end-on with a line of sight almost in the plane of the paper. All five temperature scales are in centigrade degrees, three reading upward toward the right and two toward the left. All five have been made to coincide at 40°, in whichever direction they run. In the case of the two remaining physical properties charted in figure 2, we have employed a scale of reciprocal temperatures on the Kelvin scale against the Briggs logarithms of the value of the properties, viscosity and vapor pressure, respectively.

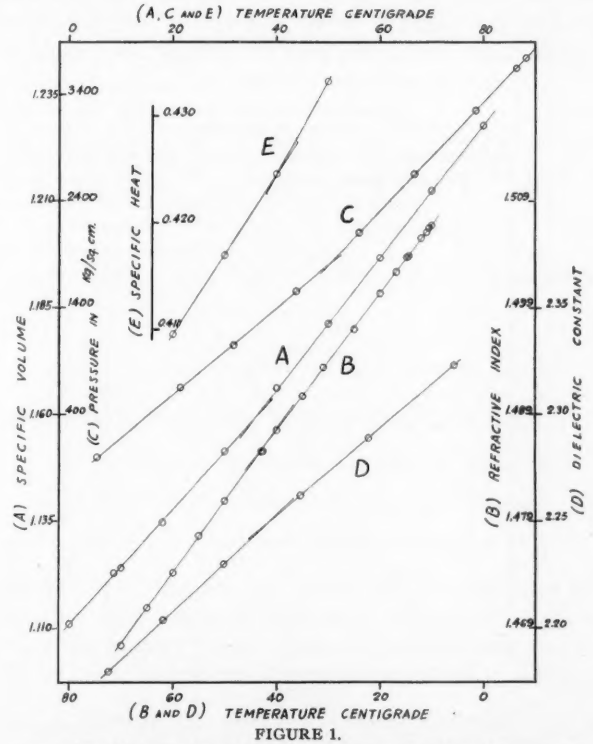



FIGURE 1.

Graphs of (A) specific volume, (B) refractive index, (C) pressure for melting, (D) dielectric constant, (E) specific heat of benzene; plotted in each case against temperature centigrade.

F. The viscosity values graphed in figure 2, curve F, are the direct observations of Thorpe and Rodger.[10] Our straight lines intersect above 30°.

G. In the case of vapor pressure, the values of the three runs of Ramsay and Young[11] were somewhat erratic, doubtless because belonging to differ-

ent runs; while the published values of Smith and Menzies[12] did not extend below 65°. We were aware that the values to be found for any physical property in International Critical Tables were likely to be smoothed values, and we, therefore, referred to the original literature for each property, so as to obtain the actual experimental values in each case. On consulting these Tables,[13] however, in order to find references to the original literature for vapor pressure, we were agreeably surprised to find that the vapor pressures of benzene below two atmospheres were reported in the form of two linear equations, in $1/T$ and $\log p$ as variables, which were stated to cover the ranges 0° to 42° and 42° to 100°, respectively. In figure 2, graph G, we have drawn these two straight lines, which intersect at 45.8°.

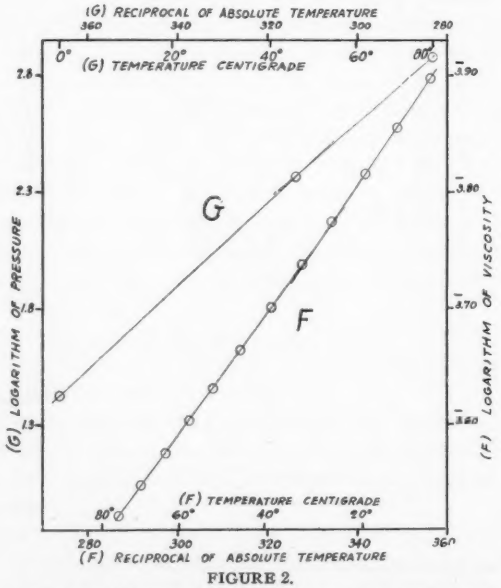



FIGURE 2.

Graphs of (F) logarithm of viscosity, (G) logarithm of vapor pressure of benzene; plotted in each against reciprocal of absolute temperature, $\times 10^5$.

It is not surprising that the breaks in the seven graphs shown above should not occur at precisely the same temperature, since the influence of a small quantity of impurity would be large on a transition of such low latent heat as is doubtless here in question.

The term allotropy is noncommittal, and has been used to cover cases both of physical and of chemical isomerism. It is applied in describing the

two forms of liquid sulphur, and may be used in describing the findings reported by Mazur[14] in the cases of liquid nitrobenzene, ethyl ether and carbon bisulphide.

In the present brief report it has been thought sufficient merely to draw attention to the phenomena observed in the case of liquid benzene. A more extended study will be published elsewhere.

[1] Young, *J. Chem. Soc.*, **55**, 504 (1889).
[2] Meyer and Mylius, *Z. physik. Chem.*, **95**, 355 (1920).
[3] Parker and Thompson, *J. Chem. Soc.*, **121**, 1341 (1922).
[4] Puschin and Grebenschtschikow, *Z. physik. Chem.*, **113**, 57 (1924).
[5] Stoops, "The Dielectric Polarization of Alcohols." Thesis deposited in Princeton University Library (1929).
[6] Smyth and Stoops, *J. Am. Chem. Soc.*, **51**, 3312 (1929).
[7] Richards and Wallace, private communication.
[8] Tréhin, *Ann. Phys.* [9], **15**, 246 (1921).
[9] Williams and Daniels, *J. Am. Chem. Soc.*, **46**, 1569 (1924).
[10] Thorpe and Rodger, *Trans. Roy. Soc. (London)*, **A185**, 522 (1894).
[11] Ramsay and Young, *Phil. Mag.* [5], **23**, 61 (1887).
[12] Smith and Menzies, *J. Am. Chem. Soc.*, **32**, 1453 (1910).
[13] *International Critical Tables*, Vol. III, p. 221.
[14] Mazur and Wolfke, *Nature*, **126**, 993 (1930); **127**, 270, 741, 893, 926; **128**, 584, 673, 761 (1931).

---

## *THE FADING OF LAKES AND DYED FIBRES**

BY WILDER D. BANCROFT AND JOHN W. ACKERMAN

BAKER CHEMICAL LABORATORY, CORNELL UNIVERSITY

Communicated December 21, 1931

A mordant is a substance that is adsorbed strongly by a textile fibre and that then takes up the dye strongly. When a dye is taken up by a mordant in the absence of a fibre, the product is called a lake. The term lake also includes dyed salts, such as barium sulphate. While it has been recognized for a long time that the fading of dyes by light in the presence of air is usually an oxidation, the data on the fading of lakes and dyes fibres are hopelessly confused.

Indigo is faster to light on wool than on cotton. One could attempt to account for this by talking about oxycellulose perhaps acting as an oxidizing agent; but then one is confronted by the statement that paranitraniline red is faster on cotton than on wool. This is in accord with the fact that methylene blue is faster on cotton than on wool; but conflicts with the observation that substantive dyes are faster on wool than on cotton. To make matters worse, paranitraniline red is theoretically a substantive dye,

though not classed as such because it is so insoluble that it cannot be applied as the regular substantive dyes are. Chemically, paranitraniline red is quite different from such a substantive dye as Congo red.

The confusion is just as bad when one comes to the dyed mordants or lakes. Quercitron, one of the natural dyes, is faster on alumina than on tin oxide, and alizarine is faster on tin oxide than on alumina, while acid green is faster on alumina than on tin oxide. Acid green is faster to light when on zinc oxide than when on tin oxide, but azo-geranine is faster on tin oxide than on zinc oxide. We are not taking into account the colored oxide of chromium, iron and copper, because the color of the mordant introduces an additional complication.

The dye solutions were made by dissolving two grams of the dye in 500 cc. of warm water and then diluting to one liter. The alumina for the alumina lakes was prepared by adding 500 cc. of a ten per cent solution of sodium carbonate rapidly at 80° with stirring to a 1200-cc. solution of a ten per cent solution of aluminum sulphate. The precipitate was washed five times and a white colloidal solution was obtained, containing by analysis about 405 milligrams $Al_2O_3$ in 50 cc. An alumina lake of alkali blue was prepared by heating 40 cc. of the dye for fifteen minutes and adding 25 cc. of the colloidal alumina solution. This was precipitated at 60° by adding 80 milligrams of iron-free aluminum sulphate dissolved in 5 cc. of water. This was then made up to a volume of 70 cc. with water and was compared before a carbon arc with a dye solution of 40 cc. alkali blue diluted to 70 cc. with water.

Glass flasks of one centimeter thickness and a capacity of 50–55 cc. were so made that a flat side would face the light. Into one flask was placed 50 cc. of the dye solution and into another 50 cc. of the suspended lake. Since the lake has a tendency to settle, a slow current of air was passed through it to keep it stirred. To make conditions comparable a similar stream of air was passed through the dye solution. Since the Fade-ometer in the laboratory was of the old type not containing an electric ventilator and humidifier, the moisture was supplied by water maintained in vessels placed on the base plate. These were kept full at all times. In order to eliminate the possible influence of any generated heat or ozone, two large electric fans were directed at the glass-enclosed carbon arc in such a manner as to drive the heat or ozone off of the top of the cabinet. Columbia violet carbons were used to form the arc.

The dye solution faded about three quarters in 75 hours' exposure. The lake had faded only slightly in that period and only a little more at the end of 164 hours. For equal weights of dye the alumina lake of alkali blue is faster to light than the straight dye. The probable explanation is that the adsorption of dye by alumina lowers the chemical potential of the dye and therefore makes it resistant to light. If this is the true explanation, then,

with equal weights of a given dye adsorbed on two different colorless mordants, that lake should be faster to light in which the dye is furthest from adsorption saturation provided there are no disturbing factors. Experiment has proved this deduction to be correct.

Indigo is adsorbed more strongly by wool than by cotton and, as we have already seen, indigo is faster to light on wool than on cotton. Acid green lakes show decreasing fastness to light in the order alumina, zinc and tin. Special experiments showed that the order of decreasing adsorption is alumina, zinc and tin. With azo-geranine the order of decreasing fastness to light is alumina, tin and zinc, and the order of decreasing adsorption is alumina, tin and zinc. With alkali blue the order of decreasing fastness to light is alumina, zinc and tin, and the order of decreasing adsorption is the same, though the difference in adsorption between hydrous zinc oxide and hydrous stannic oxide is not very great and not as much as we think it ought to be.

With alizarine the order of decreasing fastness to light is tin, alumina and zinc, while the order of decreasing adsorption is apparently alumina, tin and zinc. This is not in accord with the theory as outlined, but there are extenuating circumstances. The alumina lake is red, the tin lake is orange, and the zinc lake is lavender. While we do not yet know that this difference in color is sufficient to account for the difference in behavior to light, we can at least see that the three lakes are not strictly comparable. For the moment we shall have to leave it at that.

So far as preliminary experiments go, a given weight of sodium alizarate on stannic oxide is fully as fast as an equal weight of indigo on cotton, but not as fast as an equal weight of indigo on wool. Not too much importance can be attached to this because the two runs were not made simultaneously. We have hitherto assumed that the addition of tin salts in the dyeing with Turkey red was intended primarily to increase the brilliancy by introducing a mordant with a higher index of refraction. We now know that the tin changes the color toward orange which is apparently what people wanted and that the tin makes the color faster to light.

Since a basic mordant adsorbs much less strongly after it has aged, the dyed mordant should then be less fast to light. Special experiments have proved this to be true. Acid green was precipitated on alumina that had been aged for a month and on freshly precipitated stannic oxide. The acid green was adsorbed more strongly on the tin oxide than on the alumina and was faster to light, reversing the behavior with both mordants freshly precipitated.

While a lake pigment is ordinarily faster to light than the same weight of the dye in mass, this is not necessarily always the case. Lithopone was zinc sulphide adsorbed on barium sulphate. The early lithopones blackened readily in the light with a setting-free of metallic zinc, while zinc

sulphide will not blacken under the same conditions of exposure. The reason for this seems to be that metallic zinc is adsorbed strongly by barium sulphate, thus making it possible for light to decompose zinc sulphide.

If a basic mordant acts as an oxidizing catalyst, a dye will be less fast on it than when present in mass. There are statements in the literature than zinc oxide decreases the fastness of dyes to light. We have not checked that yet; but there is nothing impossible about it.

The general conclusions of this paper are:

1. When a colorless mordant acts solely as an adsorbing agent, lowering the chemical potential of the dye, the mordant will increase the fastness of the dye to light.

2. With equal weights of one dye on different colorless mordants, the dye on the less saturated mordant will be faster to light.

3. If the relative adsorbing powers of two colorless mordants are reversed, the fastness to light of the dyed mordants will be reversed.

4. The fastness to light of a given amount of dye on a given amount of a colorless mordant will vary with the physical state of the mordant.

5. If a mordant or a salt adsorbs one or more of the decomposition products very strongly, the lake may be less fast to light than the pure dye. No case of that is known definitely for colored lakes; but it is true for the white pigment, lithopone, as formerly made.

6. If a colorless mordant acts as an oxidizing catalyst; the dye may be less fast to light on that mordant than when pure. It is claimed that zinc oxide may act in this way.

7. An apparent exception to the relation between strength of adsorption and fastness to light was found in the alizarine lakes; but they are not the same color and therefore not strictly comparable.

* This work is done under the programme now being carried out at Cornell University and supported in part by a grant from the Heckscher Foundation for the Advancement of Research established by August Heckscher at Cornell University.

# CYTOLOGY OF THE FEROX-QUERCIFOLIA-STRAMONIUM TRIANGLE IN DATURA

BY A. DOROTHY BERGNER AND A. F. BLAKESLEE

CARNEGIE INSTITUTION OF WASHINGTON, DEPARTMENT OF GENETICS, COLD SPRING HARBOR, N. Y.

Communicated January 13, 1932

The best-known group of the genus *Datura* includes the species *D. stramonium*, *D. quercifolia* and *D. ferox*. For each of the three species an inbred line was used in making the crosses. The *stramonium* tester was our standard Line 1, which was originally obtained from Washington, D. C. The *quercifolia* tester race came from Fort Davis, Texas and the *ferox* tester from near Buenos Aires, Argentina. As has been shown earlier,[1,2] different cryptic chromosomal types occur in *D. stramonium*. In consequence different results would have been obtained if we had used certain other races of this species as a tester. In the present report we are confining ourselves to the results obtained from the use of these testers mentioned.

In the three tester races used, a cytological examination showed that the 24 chromosomes are arranged as 12 bivalents at first metaphase in pollen-mother-cells.

A similar examination of the pollen-mother-cells of the $F_1$ inter se hybrids, however, showed configurations involving two or more pairs of chromosomes as shown in figure 1. The *stramonium-quercifolia* hybrid has three configurations involving four chromosomes each plus 6 closable bivalents. Each configuration of four is a circle, or when the circle breaks, a chain. The *stramonium-ferox* hybrid also has three configurations: two circles involving four chromosomes each and a circle of four to which a bivalent is attached (configuration of six) plus five bivalents. All the circles show a tendency to break at a particular, fixed point indicated by dotted lines in the formulae. As a result they are more frequently found as chains rather than as circles. The *ferox-quercifolia* hybrid has a "figure-of-eight" configuration which involves four chromosomes plus ten bivalents. Appropriate crosses have shown that, so far as the ends of the chromosomes are concerned, the chromosomes in the two circles of four chromosomes each in the $F_1$ between Line 1 and Peruvian races of *stramonium* are the same as those found in the Line 1-*ferox* and Line 1-*quercifolia* hybrids. Four of the chromosomes of *D. ferox* and also of *D. quercifolia*, therefore, are the same, so far as their terminal attachments are concerned, as the four modified chromosomes in Peruvian races of *D. stramonium*. Throughout our discussion we shall speak of modified chromosomes in terms of the chromosomes of our standard Line 1 in *D. stramonium*.

Belling's hypothesis of segmental interchange between non-homologous chromosomes[3] will explain the occurrence of circles of four chromosomes found in intra se *stramonium* crosses between our standard Line 1 and certain races found in nature.

*Inter Se Crosses of Datura*

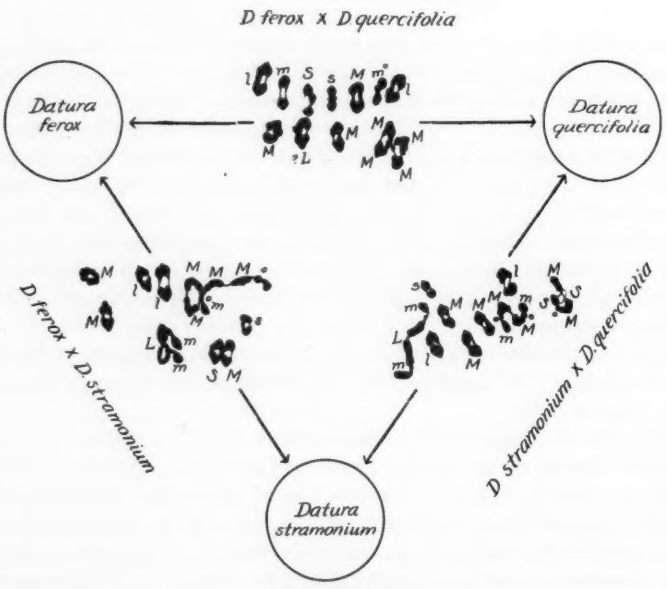
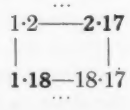



FIGURE 1

The best-known of these configurations is the B circle which is found in crosses between the *stramonium* tester and B races. Segmental interchange (reciprocal translocation) between the 1·2 and 17·18 chromosomes[4] is thought of as having taken place in the origin of the B races from Line 1 to produce the chromosomes 2·17 and 1·18. Since like ends of chromosomes attract each other, a hybrid between these two races shows a circle of four chromosomes which may be represented as follows:




In this and other diagrams in the text the numbers representing chromosomes other than the Line 1 type are shown in boldface type. Peculiari-

ties of size and shape enable one to identify the four members of the B circle. Appropriate tests have shown that both *ferox* and *quercifolia* are B races.

The other type of segmental interchange typical of *stramonium* races found in Peru involves an interchange between the $11 \cdot 12^{12}$ and $^{21}21 \cdot 22$ chromosomes.[5] The hybrid between our standard Line 1 tester and these races from Peru shows the following circle of four chromosomes represented thus:

$$\begin{array}{ccc} {}^{12}12 \cdot 11 & \text{---} & \mathbf{11 \cdot 21^{21}} \\ | & & | \\ {}^{12}\mathbf{12 \cdot 22} & \text{---} & 22.21^{21} \end{array}$$

Appropriate tests have shown that both *ferox* and *quercifolia* have this type of interchanged chromosomes.

The third configurations found in the *stramonium-quercifolia* and *stramonium-ferox* hybrids involve interchanges which have not been found within the species *stramonium* but which are related to one another and are partly identical. In the simpler form, as shown in the *stramonium-quercifolia* hybrid, this configuration is a circle of four chromosomes. In the origin of *quercifolia* we may assume that an interchange took place between the $7.8^{8}$ and the $19.20^{20}$ chromosomes of *D. stramonium* to produce the chromosomes $7.20^{20}$ and $19.8^{8}$. The *stramonium-quercifolia* hybrid then will show a circle of four chromosomes as follows:

$$\begin{array}{ccc} {}^{8}8 \cdot 7 & \text{——} & \mathbf{7 \cdot 20^{20}} \\ | & & | \\ {}^{8}\mathbf{8 \cdot 19} & \text{—} & 19 \cdot 20^{20}. \end{array}$$

The *ferox-quercifolia* hybrid showed ten bivalents plus a "figure-of-eight" configuration which involved four chromosomes. That is, insofar as the ends of the chromosomes are concerned, ten of the twelve pairs are identical in *ferox* and *quercifolia*.

The "figure-of-eight" configuration is of rather infrequent occurrence within the species *D. stramonium*. It may be interpreted as the result of mutual attraction between like humps on two bivalents; these bivalents are united in the hump region only. The explanation offered to account for this type of configuration may be thought of as a limiting case of segmental interchange between non-homologous chromosomes, namely, a reciprocal interchange involving *only* the humps which are attached to ends of the non-homologous chromosomes involved. In thus accounting for the cytological observations, we are merely extending to similar humps the theory that like ends of chromosomes are mutually attracted. Apparently like ends of two homologous chromosomes attract one another,

resulting in the formation of a closed bivalent, *regardless* of whether the humps which are attached to the two chromosomes are alike or not. If there has been an interchange of humps between two pairs of chromosomes in the origin of a particular race (or species), then in a hybrid with this race there will be two bivalents, each of which will consist of chromosomes homologous except for their attached humps. The mutual attraction between the two pairs of identical humps which are now located on different bivalents results in the joining together of these bivalents by the humps only.

Thus we can account for the "figure-of-eight" configuration which is found in the *ferox-quercifolia* hybrid. By an interchange of humps between the $7{\cdot}20^{20}$ and $15{\cdot}16^{16}$ chromosomes which we may assume to have taken place in the origin of *D. ferox* from *D. quercifolia*, there resulted the $7{\cdot}20^{16}$ and $15{\cdot}16^{20}$ chromosomes. The *ferox-quercifolia* hybrid then would show a "figure-of-eight" configuration of four chromosomes with the following composition:

$$\begin{array}{c} 7{\cdot}20^{20}\text{—}^{20}16{\cdot}15 \\ \langle \quad \rangle \quad \langle \quad \rangle \\ 7{\cdot}20^{16}\text{—}^{16}16{\cdot}15. \end{array}$$

The mutual attraction between like humps results in the attachment between these bivalents at their humps only.

The third configuration in the *stramonium-ferox* hybrid differs from the third configuration in the *stramonium-quercifolia* hybrid in that it involves six chromosomes. It is a circle of four to which a bivalent is attached by its humps. In analyzing the "figure-of-eight" configuration in the *ferox-quercifolia* hybrid above, we have already designated two of the *ferox* chromosomes as $7{\cdot}20^{16}$ and $15{\cdot}16^{20}$. Another *ferox* chromosome involved in this third configuration must be $19{\cdot}8^{8}$, as will be seen from the following consideration. The sizes and connections of chromosomes in the *stramonium-quercifolia* third configuration showed that one of the *quercifolia* chromosomes must be designated as $19{\cdot}8^{8}$. This chromosome in the *ferox-quercifolia* hybrid forms a closed bivalent and therefore *D. ferox* must have the same chromosome ($19{\cdot}8^{8}$) so far as its ends are concerned. In the *stramonium-ferox* hybrid, then, the configuration of 6 chromosomes (a circle to which a bivalent is attached) may be explained as follows:

$$\begin{array}{l} {}^{8}8{\cdot}7\text{——}7{\cdot}20^{16}\text{—}^{16}16{\cdot}15 \\ \quad | \qquad\quad | \quad \langle \quad \rangle \\ {}^{8}8{\cdot}19\text{—}19{\cdot}20^{20}\text{—}^{20}16{\cdot}15. \end{array}$$

This third configuration in the *stramonium-ferox* hybrid has proved interesting because of its instability in back-crosses to *stramonium*. A

few plants of two back-crosses were examined cytologically with the results shown in table 1. The configuration of six chromosomes was recovered only once in thirty plants. Eight plants, however, showed a circle of four chromosomes which appears to be identical with the third

TABLE 1
FEROX-STRAMONIUM $F_1$ BACK-CROSS

| No. plants examined | Pedigree number | B (L+m) | Peruvian ($M^{\circ}+S$) | Ferox 3rd ($M^{\circ}+m^{\circ}+M^{\circ}$) | Extracted from 3rd: 3a ($M^{\circ}+m^{\circ}$) | Extracted from 3rd: 3b ($M^{\circ}+m^{\circ}$) | 12 bivalents | Open L bv. |
|---|---|---|---|---|---|---|---|---|
| 20 | 290382 | 15 | 5 | 0 | 6 | 2 | 1 | 1 |
| 10 | 3001200 | 5 | 3 | 1 | 2 | 2 | 1 | 0 |
| 30 | | 20 | 8 | 1 | 8 | 4 | 2 | 1 |

configuration found in the *stramonium-quercifolia* hybrid and its back-crosses. In addition four other plants showed a "figure-of-eight" configuration which involved four chromosomes. This could not be identical with the "figure-of-eight" configuration that was found in the *ferox-quercifolia* hybrid because of a difference in the sizes of the chromosomes involved. In the *ferox-quercifolia* hybrid the two connected bivalents were M° and M°; in the *stramonium-ferox* back-cross they were M° and m°. Both of these types have been recovered in the second and third back-crosses. The resolution of the configuration of six chromosomes into two separate configurations of four each needs explanation. From cytological evidence it appears that crossing-over between homologous parts of chromosomes will account for this change in configurations. Our hypothesis of how this has taken place will be made clearer from the diagram in figure 2. We assume that crossing-over in the *stramonium-ferox* hybrid took place between the $7{\cdot}20^{16}$ and the $19{\cdot}20^{20}$ chromosomes somewhere in the homologous region below the humps. Such crossing-over between homologous regions of otherwise non-homologous chromosomes has been demonstrated by a number of investigators[6] in such diverse forms as *Drosophila*, *Zea mays*, *Pisum* and *Oenothera*. The newly formed chromosomes then may be designated as $7{\cdot}20^{20}$ and $19{\cdot}20^{16}$. No evidence is available as yet regarding the exact place at which the crossing-over has taken place. The cytological results observed would be satisfied if the crossing-over involved only an interchange of the humps.

Ordinarily the result of such crossing-over is not detectable cytologically because the interchanged chromosomes consist of regions derived from only two chromosomes. In the *stramonium-ferox* hybrid the situa-

tion is different. In terms of the *stramonium* tester, the $7{\cdot}20^{16}$ chromosome consists of parts of three chromosomes: its 7 region is homologous with the 7 region of the $7{\cdot}8^{8}$ chromosome, the 20 region with the 20 region of the $19{\cdot}20^{20}$ chromosome, and the 16 hump with the 16 hump of the $15{\cdot}16^{16}$ chromosome. Therefore, if crossing-over takes place in the 20 region of the $7{\cdot}20^{16}$ and $19{\cdot}20^{20}$ chromosomes, the resulting chromosomes will be $7{\cdot}20^{20}$ and $19{\cdot}20^{16}$. Such crossing-over is possible *only* when the organism is heterozygous for the interchanged chromosomes. This is the situation in the *stramonium-ferox* hybrid.

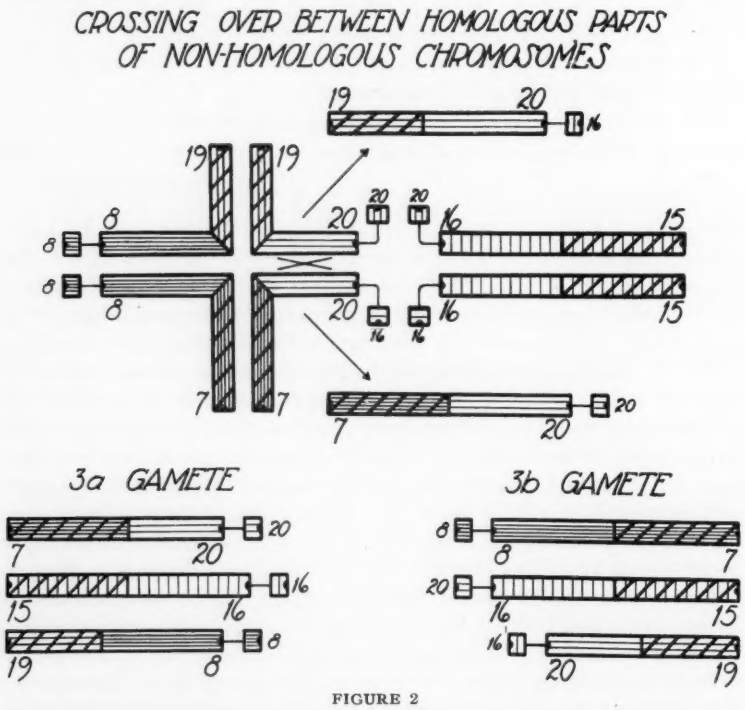



FIGURE 2

This heterozygous condition also makes possible the incorporation of the newly formed chromosomes into viable gametes. The $7{\cdot}20^{20}$ chromosome if combined with the $19{\cdot}8^{8}$ and $15{\cdot}16^{16}$ chromosomes will form a viable gamete; likewise the $19{\cdot}20^{16}$ chromosome if combined with the $15{\cdot}16^{20}$ and $7{\cdot}8^{8}$ chromosomes will form a viable gamete. If each of these gametes is united with a gamete of the *stramonium* tester, as is the case in back-crosses to the *stramonium* tester, then two kinds of configurations will result. The circle of four will be:

$$\begin{array}{ccc} {}^{8}8\cdot 7 & \text{——} & \mathbf{7\cdot 20}^{20} \\ | & & | \\ {}^{8}\mathbf{8\cdot 19} & \text{—} & 19\cdot 20^{20} \end{array}$$

The "figure-of-eight" configuration will be:

$$\begin{array}{ccccc} 15\cdot 16^{16} & \text{—} & {}^{16}\mathbf{20\cdot 19} & & \\ \langle & \rangle & \langle & \rangle & \\ \mathbf{15\cdot 16}^{20} & \text{—} & {}^{20}20\cdot 19 & & \end{array}$$

Cytologically the circle of four is indistinguishable from the third configuration found in the *stramonium-quercifolium* hybrid. Certainly as far as the ends of the involved chromosomes are concerned it is identical with it. The "figure-of-eight" configuration is new and promises to be a useful tester.

It is now possible to summarize the analysis of the chromosomes involved in the configurations found in these interspecific hybrids. *D. ferox* has five chromosomes, out of a possible twelve, whose ends are identical with the homologous chromosomes in the *stramonium* tester; *D. quercifolia* has six. The numerical representation of the seven *ferox* and six *quercifolia* chromosomes which d.ffer from those in the *stramonium* tester are listed in table 2. In this table also are listed the two new chromosomes extracted from the third configuration in the *stramonium-ferox* hybrid.

TABLE 2

CHROMOSOMES INVOLVED IN THE CONFIGURATIONS FOUND IN STRAMONIUM-QUERCIFOLIA-FEROX HYBRIDS

| DATURA STRAMONIUM | | D. QUERCIFOLIA | D. FEROX | | |
|---|---|---|---|---|---|
| Rolled (L) | 1.2 | 2.17 | 2.17 | | |
| Poinsettia (m) | 17.18 | 1.18 | 1.18 | | |
| Cocklebur (M°) | $11.12^{12}$ | $11.21^{21}$ | $11.21^{21}$ | EXTRACTED FEROX TYPES | |
| Globe (S°) | $^{21}21.22$ | $^{12}12.22$ | $^{12}12.22$ | *3a* | *3b* |
| Elongate (M°) | $7.8^{8}$ | $7.20^{20}$ | $7.20^{16}$ | $7.20^{20}$ | $7.8^{8}$ |
| Spinach (m°) | $19.20^{20}$ | $19.8^{8}$ | $19.8^{8}$ | $19.8^{8}$ | $19.20^{16}$ |
| Reduced (M°) | $15.16^{16}$ | $15.16^{16}$ | $15.16^{20}$ | $15.16^{16}$ | $15.16^{20}$ |

The letters in connection with the chromosomes indicate their relative sizes: L signifies very large, M, medium, m, small medium and S, small. The circle above some of the letters as well as the raised numbers represent humps on these chromosomes.

The isolation of each of these separate "Prime Types" in the homozygous condition is in the process of being made. It is planned ultimately to recombine these "Prime Types" with otherwise Line 1 *stramonium* tester chromosomes to form the resynthesized species *ferox* and *quercifolia*.

As far as the ends of the chromosomes are concerned, *D. ferox* and *D. quercifolia* are closely related.

It should be pointed out again that in our discussions we have classified chromosomal evolution in terms of the chromosomes of our *stramonium* tester Line 1. This is merely a matter of convenience since we are most familiar with the chromosomes of this particular race. It is possible if not probable, that our tester race is a derived rather than a primitive type as may be seen from the following consideration: An examination of more than 500 natural races of *D. stramonium* indicates that the B type is the predominant one except in the United States and Brazil. Our *stramonium* tester Line 1 therefore may equally well represent a segmental interchange that has occurred fairly recently in the evolution of the *stramonium* species. Furthermore, it is known that the *ferox* and *quercifolia* as well as the *leichardtii* tester races are B types. But the third types of interchange which are common to *ferox* and *quercifolia* have not been found in any *stramonium* races.

It should be emphasized that the chromosomal constitutions that have been described refer only to the particular testers used as representatives of *D. stramonium*, *D. ferox* and *D. quercifolia*. It is known that certain other races of *D. stramonium* would have given different results because of segmental interchanges which apparently took place in their origin. Such cryptic types have been demonstrated by intra se crosses. Although *D. ferox* and *D. quercifolia* have not been studied as yet from this standpoint, these species presumably also contain cryptic types. In an examination of ten races of *D. meteloides* crossed to a *meteloides* tester, two configurations due to segmental interchange were identified in the $F_1$ with one race. Likewise of three races of *D. innoxia* which were crossed to an *innoxia* tester, one induced a configuration due to interchange. In still another species, *D. metel*, one of the two races which were crossed to a *metel* tester induced a single and the other induced two separate configurations. Therefore the presence of cryptic types has been demonstrated in four species of Datura; in fact, in all the species of this genus that so far have been investigated from this standpoint. It is probable, therefore, that cryptic chromosomal types due to segmental interchange are of wide-spread occurrence.

So far we have found chromosomal configurations in eight different inter se hybrids which involved seven species of *Datura*. In all of these we have found configurations interpreted as due to segmental interchange. Segmental interchange, therefore, may be assumed to have taken place in the differentiation of these species. Just what rôle, however, this process has played in speciation is not yet clear.

*Summary.*—Three inbred tester races were used in making inter specific hybrids involving *Datura stramonium*, *D. quercifolia* and *D. ferox*. The results which have been obtained apply to these testers only.

Both *D. ferox* and *D. quercifolia* are "B" and "Peruvian" types. They

show, in common with races of *stramonium* from Peru, two types of segmental interchange between non-homologous chromosomes. The B race has the *stramonium* tester chromosomes 1·2 and 17·18 modified to 2·17 and 1·8. The Peruvian type has the *stramonium* tester chromosomes $11{\cdot}12^{12}$ and $^{21}21{\cdot}22$ modified to $11{\cdot}21^{21}$ and $^{12}12{\cdot}22$.

A previously unknown type of interchange is partly identical in *ferox* and *quercifolia*. In *quercifolia*, the *stramonium* tester chromosomes $7{\cdot}8^{8}$ and $19{\cdot}20^{20}$ are modified to $7{\cdot}20^{20}$ and $19{\cdot}8^{8}$. In *ferox* a further interchange between the $7{\cdot}20^{20}$ and $15{\cdot}16^{16}$ chromosomes has formed the $7{\cdot}20^{16}$ and $15{\cdot}16^{20}$ chromosomes.

In back-crosses of the *stramonium-ferox* hybrid this configuration has given rise to two new configurations, both of which are explainable on the assumption of crossing-over between the homologous regions of the $7{\cdot}20^{16}$ and $19{\cdot}20^{20}$ chromosomes. Such crossing over is detectable in the hybrid because in terms of the *stramonium* tester the $7{\cdot}20^{16}$ chromosome is composed of parts of these three chromosomes $7{\cdot}8^{8}$, $15{\cdot}16^{16}$ and $19{\cdot}20^{20}$. The heterozygous condition of the hybrid makes possible the incorporation of the newly formed chromosomes into viable gametes.

So far as investigated, these species of *Datura* show differences in the architecture of their chromosomes which has been interpreted as due to segmental interchange.

## BIBLIOGRAPHY AND NOTES

[1] Blakeslee, A. F., *J. Hered.*, **20**, 177–190 (1929).

[2] Bergner, A. D., and Blakeslee, A. F., *5th Intern. Bot. Cong., Report of Proceedings (Sect. Abst.)*, **126** (1930).

[3] Belling, J., and Blakeslee, A. F., these PROCEEDINGS, **12**, 7–11 (1926).

[4] The names and relative sizes of the chromosomes, which are here given by their numbered ends, are shown in table 2. The use of numbered ends to designate particular chromosomes has been employed previously in joint papers by A. F. Blakeslee and R. E. Cleland, these PROCEEDINGS, **16**, 177–189 (1930).

[5] The superscripts to the ends 12 and 21 represent terminal humps characteristic of these Line 1 chromosomes. G. A. Lewitsky has demonstrated satellites or knobs on several of the chromosomes in root-tips of *D. stramonium*, *Proc. U. S. S. R. Cong. Genet.*, **2**, 87–105 (1929). They are probably the equivalents of those humps found on meiotic chromosomes. Such humps are convenient markers to identify chromosomes and particular ends of chromosomes.

[6] Brink, R. A., and Cooper, D. C., *Genet.*, **16**, 595–628 (1931). Burnham, C. R., these PROCEEDINGS, **16**, 269–277 (1930). Creighton, H. B., and McClintock, B., these PROCEEDINGS, **17**, 485–497 (1931). Dobzhansky, Th., *Amer. Nat.*, **65**, 214–232 (1931). Emerson, S., and Sturtevant, A. H., *Zeitschr. indukt. Abst.-u. Vererb.*, **59**, 395–419 (1931). Håkansson, A., *Hereditas*, **15**, 2–61 (1931). Muller, H. J., *J. Genet.*, **22**, 335–358 (1930). Rhoades, M. M., *Genet.*, **16**, 490–504 (1931).

# A POSSIBLE INFLUENCE OF THE SPINDLE FIBRE ON CROSSING-OVER IN DROSOPHILA

BY G. W. BEADLE[1]

WILLIAM G. KERCKHOFF LABORATORIES OF THE BIOLOGICAL SCIENCES, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA

Communicated December 19, 1931

The distribution of genes in the second and third chromosomes of *Drosophila melanogaster* suggests that crossing-over may be influenced by the spindle fibre at least in neighboring regions. The disparity between the genetic and cytological maps of these chromosomes especially in the spindle fibre regions (Dobzhanzky[2,3]) is consistent with such a view. Sturtevant[4] suggests from studies of crossing-over in Drosophila stocks homozygous for inverted sections of the autosomes that crossing-over in a given region is a function of the material in that region rather than of its position in the chromosome. He found no significant changes in amount of crossing-over with change in position of a section of a chromosome with respect to the spindle fibre or the end of the chromosome. Muller and Stone[5] in a study of crossing-over in the case of a homozygous inversion in the X-chromosome of Drosophila apparently found that the regions in the inverted section gave normal crossing-over. However, since in none of these cases of homozygous inversions did the inversion include regions close to the spindle fibre, the possibility remains that the spindle fibre influences crossing-over in regions close to it.

Dobzhansky[6] studied crossing-over in a homozygous III–IV translocation in Drosophila (designated as translocation-*c*) and found no significant differences from normal controls. However, the regions which might have been expected to be influenced by the spindle fibre were not studied. At the suggestion of Professor Dobzhansky, the writer has made additional studies of crossing-over in flies homozygous for this translocation. The results are reported in the present paper.

In the translocation stock under consideration a distal segment made up of about three-fourths of the right limb of chromosome III (cytologically[3]) is broken off and attached to chromosome IV The break in chromosome III is close to and to the left of the curled gene (0.7 unit according to Dobzhanky's data[3] and somewhat closer according to results obtained by the writer). Accordingly, in the homozygous translocation flies the section of the third chromosome including curled and regions to the right is closer to a spindle fibre (of the fourth chromosome) by a cytological distance equal to about one-fourth the length of the right limb of chromosome III. Genetically this distance represents about 2 crossover units (standard map[7]).

A stock of the homozygous translocation was made up carrying in one of the translocated third chromosomes the genes roughoid (*ru*), hairy (*h*), thread (*th*), scarlet (*st*), curled (*cu*), stripe (*sr*), sooty ($e^s$) and claret (*ca*) and in the homolog the normal allelomorphs of these genes. The constitution of this stock was therefore $\frac{+\ +\ +\ +}{ru\ h\ th\ st}\ \frac{+\ +\ +\ +}{cu\ sr\ e^s\ ca}$. All of the recessives were introduced, by crossing-over, from a normal stock. The normal allelomorphs of *ru*, *h*, *th*, *st*, $e^s$ and *ca* were likewise introduced from a normal stock. Females of the above-mentioned constitution were used for a study of crossing-over by mating with homozygous translocation males carrying the above-named recessives in both third chromosomes. The control was made up by using the stocks which had been used in introducing the desired genes into the translocation stock.

The standard map[7] positions of the loci used are shown in figure 1.

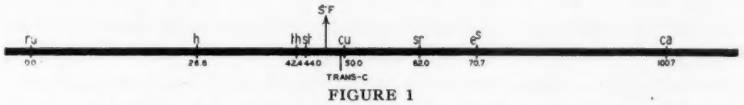



FIGURE 1

Standard map positions of the loci used. S F refers to the spindle fibre.

Crossover data from the homozygous translocation stock and the control are presented in table 1. These data together with the crossover

TABLE 1

CROSSOVER DATA FROM HOMOZYGOUS TRANSLOCATION-*c* FEMALES

$$\left(\frac{ru\ \ h\ \ th\ \ st\ \ cu\ \ sr\ \ e^s\ \ ca}{1\ \ 2\ \ 3\ \ 4\ \ 5\ \ 6\ \ 7} \times ru\ \ h\ \ th\ \ st\ \ cu\ \ sr\ \ e^s\ \ ca\right)$$

AND A NORMAL CONTROL OF SIMILAR CONSTITUTION

(The class carrying the dominant allelomorph of *ru* is given first in each case)

| CROSSOVERS REGION | TRANSLOCATION | CONTROL | CROSSOVERS REGION | TRANSLOCATION | CONTROL |
|---|---|---|---|---|---|
| 0 | 519 | 140 | 2, 3 | 0 | 1 |
| | 330 | 107 | | 0 | 0 |
| 1 | 199 | 50 | 2, 4 | 168 | 1 |
| | 255 | 57 | | 131 | 0 |
| 2 | 140 | 36 | 2, 5 | 2 | 6 |
| | 151 | 45 | | 3 | 6 |
| 3 | 3 | 1 | 2, 6 | 3 | 8 |
| | 3 | 0 | | 6 | 5 |
| 4 | 450 | 8 | 2, 7 | 79 | 25 |
| | 442 | 8 | | 84 | 27 |
| 5 | 6 | 25 | 3, 4 | 4 | 0 |
| | 3 | 20 | | 0 | 0 |
| 6 | 29 | 34 | 3, 5 | 1 | 0 |
| | 21 | 22 | | 0 | 0 |

TABLE 1 (*Continued*)

| CROSSOVERS REGION | TRANS-LOCATION | CONTROL | CROSSOVERS REGION | TRANS-LOCATION | CONTROL |
|---|---|---|---|---|---|
| 7 | 269 | 85 | 3, 6 | 2 | 0 |
| | 213 | 54 | | 0 | 0 |
| 1, 2 | 45 | 12 | 3, 7 | 5 | 0 |
| | 25 | 7 | | 7 | 1 |
| 1, 3 | 1 | 0 | 4, 5 | 6 | 1 |
| | 2 | 1 | | 4 | 1 |
| 1, 4 | 240 | 2 | 4, 6 | 10 | 3 |
| | 199 | 5 | | 25 | 2 |
| 1, 5 | 1 | 6 | 4, 7 | 200 | 8 |
| | 3 | 9 | | 240 | 5 |
| 1, 6 | 8 | 9 | 5, 6 | 0 | 0 |
| | 7 | 9 | | 0 | 1 |
| 1, 7 | 83 | 33 | 5, 7 | 1 | 12 |
| | 124 | 36 | | 1 | 10 |
| 6, 7 | 1 | 3 | 2, 4, 7 | 82 | 2 |
| | 0 | 4 | | 77 | 1 |
| 1, 2, 4 | 19 | 1 | 2, 5, 7 | 0 | 5 |
| | 30 | 1 | | 0 | 8 |
| 1, 2, 5 | 0 | 1 | 2, 6, 7 | 0 | 0 |
| | 0 | 2 | | 3 | 0 |
| 1, 2, 6 | 0 | 0 | 3, 4, 7 | 3 | 0 |
| | 1 | 3 | | 7 | 0 |
| 1, 2, 7 | 13 | 4 | 4, 5, 7 | 2 | 0 |
| | 8 | 4 | | 0 | 0 |
| 1, 3, 4 | 4 | 0 | 4, 6, 7 | 0 | 0 |
| | 1 | 0 | | 1 | 1 |
| 1, 4, 5 | 2 | 0 | 5, 6, 7 | 0 | 0 |
| | 4 | 0 | | 0 | 2 |
| 1, 4, 6 | 8 | 0 | 1, 2, 4, 7 | 13 | 0 |
| | 7 | 1 | | 16 | 0 |
| 1, 4, 7 | 118 | 1 | 1, 2, 5, 7 | 0 | 1 |
| | 102 | 4 | | 0 | 0 |
| 1, 5, 6 | 0 | 1 | 1, 3, 4, 7 | 0 | 0 |
| | 0 | 0 | | 1 | 0 |
| 1, 5, 7 | 0 | 3 | 1, 4, 5, 7 | 0 | 0 |
| | 0 | 7 | | 1 | 1 |
| 1, 6, 7 | 0 | 3 | 1, 4, 6, 7 | 2 | 0 |
| | 0 | 6 | | 0 | 0 |
| 2, 4, 5 | 1 | 1 | 2, 4, 5, 7 | 0 | 1 |
| | 3 | 2 | | 1 | 0 |
| 2, 4, 6 | 6 | 0 | 2, 4, 6, 7 | 2 | 0 |
| | 9 | 1 | | 0 | 0 |
| | | | 1, 2, 4, 5, 7 | 0 | 0 |
| | | | | 1 | 0 |
| | | | Total | 5303 | 1019 |

data from the heterozygous translocation flies (from Dobzhansky[5]) are summarized in table 2. As compared with the control, crossing-over in

TABLE 2

SUMMARY OF CROSSOVER DATA FROM HOMOZYGOUS TRANSLOCATION-*c*, HETEROZYGOUS TRANSLOCATION-*c* (FROM DOBZHANSKY[5]) AND NORMAL FLIES

| EXPERIMENT | PER CENT CROSSING-OVER IN REGION 1 | 2 | 3 | 4 | 5 | 6 | 7 | NO. INDIVIDUALS |
|---|---|---|---|---|---|---|---|---|
| Homozygous translocation | 29.1 | 21.2 | 0.8 | 49.9 | 0.9 | 2.8 | 33.2 | 5303 |
| Heterozygous translocation | 26.9 | 16.9 | 0.9 | 4.5 | 3.7 | 7.5 | 28.9 | 3185 |
| Control | 27.5 | 21.3 | 0.4 | 6.1 | 13.3 | 11.6 | 35.0 | 1019 |

flies carrying the translocation in homozygous form is reduced in the *cu-sr* interval from 13.3 to 0.9 and in the *sr-e*[8] interval from 11.6 to 2.8 per cent. Other comparable regions show no differences. Crossing-over in flies heterozygous for the translocation is also decreased as compared with the controls in the *cu-sr* and *sr-e*[8] intervals. This decrease has been explained by Dobzhansky[8] on the basis of decreased frequency of synapsis. The fourth chromosome to which a section of chromosome III is attached presumably exerts a mutual attraction with its free homolog. Likewise, the two sections of chromosome III and the homologous unbroken third chromosome are attracted together. These counter-forces working on the broken third chromosome would be expected to interfere with synapsis and presumably also with crossing-over. This explanation obviously cannot be applied to the homozygous translocation form. One must, therefore, seek another explanation of the decrease.

Crosses were made to test the possibility that the observed decrease in crossing-over might have been due to the presence of an extra fourth chromosome. It is known from Dobzhansky's cytological work[3] on the original stock of the homozygous translocation that an extra fourth chromosome is probably not necessary for the survival and fertility of flies homozygous for the translocation. Crosses of homozygous translocation flies with flies carrying dominant eyeless in one-fourth chromosome and minute-IV in the other gave only two classes, eyeless and minute-IV. A total of 22 flies homozygous for the translocation were so tested. Since two normal allelomorphs suppress minute-IV this test indicates that the homozygous translocation flies did not carry an extra fourth chromosome. Two normal allelomorphs also suppress dominant eyeless but since the extra fourth chromosome would have carried the recessive eyeless-2 gene this test may not be valid. Crosses of 11 homozygous translocation flies with haplo-IV gave normal and haplo-IV flies in varying ratios but in two cases these were such (215 to 161 and 105 to 116) that the translocation parent could hardly have carried an extra fourth chromosome. The fact that haplo-IV flies carrying the translocation survive

and are fertile shows that the fourth chromosome to which the section of chromosome III has been translocated does not carry a lethal. There is therefore no apparent reason why an extra fourth chromosome should be necessary to flies carrying the translocation in homozygous form. The above tests make it very improbable that the homozygous translocation stocks carried an extra fourth chromosome or a fragment of one which might have accounted for the observed decrease in crossing-over.

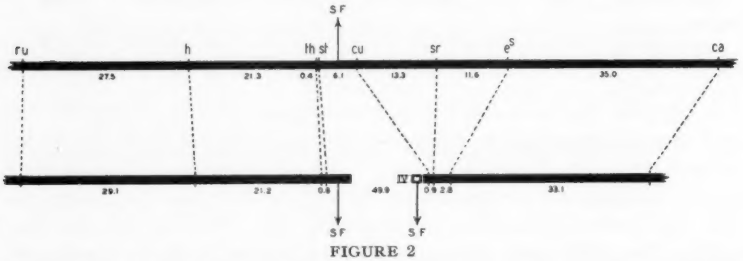



FIGURE 2

Comparison of crossing-over in control and in homozygous translocation-*c*. Figures represent observed percentages of crossing-over. S F refers to spindle fibres. The fourth chromosome is designated IV.

The above data indicate that the decrease in crossing-over in the *cu-sr* and *sr-e*$^s$ intervals in the translocation stock is due to these intervals being closer to the spindle fibre (of the fourth chromosome) than in the control by a distance equal approximately to one-fourth the length (cytologically) of the right limb of chromosome III.

On the basis of the hypothesis of crossing-over recently presented by Sax[9] the spindle fibre should reduce crossing-over in its immediate neighborhood. If the opening out of chromatids at diplotene is determined in the reductional plane at the spindle fibre then interference with crossing-over for a distance on both sides of the spindle fibre would be expected. On the basis of the crossing-over hypothesis put forward by Janssens[10] and recently supported by Darlington[11] or on the scheme of Belling[12] there is no apparent reason why the spindle fibre should interfere with crossing-over.

Coincidence values across the break in chromosome III have been calculated. They are as follows:

| REGIONS | COINCIDENCE |
|---|---|
| 1–6 | 0.76 |
| 2–6 | 0.90 |
| 1–7 | 0.94 |
| 2–7 | 1.01 |

As might have been expected there is no significant interference of crossing-over between the two independent sections of chromosome III.

*Summary.*—Crossing-over data from a homozygous III–IV translocation in *Drosphila melanogaster* indicate that the spindle fibre interferes with crossing-over in its immediate neighborhood.

*Acknowledgments.*—The writer is grateful to Professor T. Dobzhansky who suggested the study here reported and furnished the original stock carrying the translocation studied. Professor A. H. Sturtevant has made many helpful suggestions for which the writer is thankful.

[1] NATIONAL RESEARCH COUNCIL FELLOW in the Biological Sciences.

[2] Dobzhansky, T., *Genetics*, **16**, 629–658 (1931).
[3] Dobzhansky, T., *Ibid.*, **15**, 347–399 (1930).
[4] Sturtevant, A. H., *Carnegie Inst. Wash. Pub.*, **421**, 1–27 (1931).
[5] Muller, H. J., and W. S. Stone, *Anat. Rec.*, **47**, 393–394 (1930).
[6] Dobzhansky, T., *Proc. Nat. Acad. Sci.*, **15**, 633–638 (1929).
[7] Morgan, T. H., C. B. Bridges, and A. H. Sturtevant, *Bib. Genet.*, **2**, 1–262 (1925).
[8] Dobzhansky, T., *Amer. Nat.*, **65**, 214–232 (1931).
[9] Sax, Karl, *Journ. Arnold Arboretum*, **11**, 193–220 (1930).
[10] Janssens, F. A., *La Cellule*, **25**, 387–411 (1909).
[11] Darlington, C. D., *Proc. Roy. Soc.*, **B107**, 50–59 (1930).
[12] Belling, J., *Univ. Calif. Pub. Bot.*, **16**, 311–338 (1931).

---

## *AN EXPANSION OF MEROMORPHIC FUNCTIONS*

BY J. L. WALSH

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated January 9, 1932

It is the primary object of this note to exhibit a development in series of an arbitrary meromorphic function, a development which represents the function throughout its entire domain of definition, and which can be chosen in an infinite variety of ways.

THEOREM I. *If $f(z)$ is a meromorphic function of $z$ (that is, analytic except possibly for poles, at every finite point of the plane), if all the poles of $f(z)$ belong to the sequence $\alpha_1, \alpha_2, \alpha_3, \ldots \longrightarrow \infty$, where every pole occurs a number of times at least corresponding to its multiplicity, and if the numbers $\beta_1, \beta_2, \beta_3, \ldots$ are distinct from the $\alpha_n$ and uniformly limited, then there exists a unique expansion*

$$f(z) = a_0 + a_1 \frac{z - \beta_1}{z - \alpha_1} + a_2 \frac{(z - \beta_1)(z - \beta_2)}{(z - \alpha_1)(z - \alpha_2)} + \ldots \qquad (1)$$

*valid for all finite values of $z$ other than the points $\alpha_i$. The series (1) converges absolutely for all values of $z$ other than the $\alpha_i$ and uniformly in any*

*finite closed region containing no point $\alpha_i$. In such a region the convergence of* (1) *is more rapid than that of any geometric series.*

*In particular we may choose $\beta_1 = 0$, $\beta_{m+1} = 1/\bar{\alpha}_m$. Then if $f(z)$ is analytic for $|z| \leqq 1$ and if $|\alpha_m| > 1$, the sum of the first $n + 1$ terms of* (1) *is the rational function of degree $n$ whose poles lie in the points $\alpha_1, \alpha_2, \ldots, \alpha_n$, of best approximation to $f(z)$ on C: $|z| = 1$ in the sense of least squares.*

*If $f(z)$ is analytic at $z = 0$ and we choose $\beta_i = 0$, $\alpha_i \neq 0$, then the sum of the first $n + 1$ terms of* (1) *is the rational function of degree $n$ whose poles lie in the points $\alpha_1, \alpha_2, \ldots, \alpha_n$ which has contact of highest order with $f(z)$ at the point $z = 0$.*

Similarly, if we choose $\beta_1 = 0$, $\beta_{m+1} = r^2/\bar{\alpha}_m$, $|\alpha_m| > r$, and if $f(z)$ is analytic for $|z| \leqq r$, then the sum of the first $n + 1$ terms of (1) is the rational function of degree $n$ whose poles lie in the points $\alpha_1, \alpha_2, \ldots, \alpha_n$, of best approximation to $f(z)$ on $|z| = r$ in the sense of least squares.

If the function $f(z)$ is an entire function, the development (1) is valid for all values of $z$ or for all values of $z$ except the $\alpha_n$, according as we have $\alpha_n = \infty$ or $\alpha_n \longrightarrow \infty$. In the former case, $\beta_n = 0$ gives both best approximation in the sense of least squares on an arbitrary circle $|z| = r$, and contact of highest order at the origin. If we have $\alpha_n = \infty$, $\beta_n = 0$, the expansion (1) is the Taylor expansion of $f(z)$ about the origin.

Even though the values $\alpha_n = \infty$ and $\beta_n = \infty$ cannot be used in (1), according to a literal interpretation of the formula, we interpret that formula as admitting the use of those values—although in Theorem I itself the value $\beta_n = \infty$ cannot occur. Moreover, we do not suppose the $\alpha_n$ or the $\beta_n$ all distinct, but suppose merely that no $\alpha_n$ shall equal a $\beta_i$.

If the $\beta_n$ are all distinct, an expansion (1) of a given function valid in the points $\beta_n$ is unique. For if we set $z = \beta_1$ we have the condition

$$a_0 = f(\beta_1).$$

If we now set $z = \beta_2$, we have

$$f(\beta_2) = a_0 + a_1 \frac{\beta_2 - \beta_1}{\beta_2 - \alpha_1}$$

so that $a_1$ is determined. In a similar manner we see that all the coefficients $a_n$ are successively uniquely determined. This same fact (uniqueness of the determination of the coefficients $a_n$ as a result of the validity of (1) in the points $\beta_n$) also holds even if the $\beta_n$ are not all distinct, provided we make the convention that the validity of (1) at a point $\beta$ at which $k$ of the points $\beta_n$ coincide shall be taken as equivalent to the representation of $f(\beta)$, $f'(\beta)$, $\ldots$, $f^{(k-1)}(\beta)$ by (1) and the respective derived series of (1). Under this convention, the coefficient $a_n$ is determined as follows. Let $a_0, a_1, \ldots, a_{n-1}$ be known and let precisely $k$ of the numbers $\beta_1, \beta_2, \ldots, \beta_n$ coincide with $\beta_{n+1}$. If we differentiate (1) formally $k$ times

and set $z = \beta_{n+1}$, we see that $a_n$ is uniquely determined by the resulting equation.

Another way of expressing the formal relationship between $f(z)$ and the $a_n$, by assuming (1) valid in the points $\beta_n$, is to determine $a_n$ or $a_{n+k}$ after $a_0, a_1, \ldots, a_{n-1}$ have been determined, by the relation

$$\frac{(z-\alpha_1)(z-\alpha_2)\ldots(z-\alpha_n)}{(z-\beta_1)(z-\beta_2)\ldots(z-\beta_n)}\left[f(z) - a_0 - a_1\frac{z-\beta_1}{z-\alpha_1} - \ldots - a_{n-1}\frac{(z-\beta_1)(z-\beta_2)\ldots(z-\beta_{n-1})}{(z-\alpha_1)(z-\alpha_2)\ldots(z-\alpha_{n-1})}\right] = a_n + a_{n+1}\frac{z-\beta_{n+1}}{z-\alpha_{n+1}} + a_{n+2}\frac{(z-\beta_{n+1})(z-\beta_{n+2})}{(z-\alpha_{n+1})(z-\alpha_{n+2})} + \ldots, \qquad (2)$$

and this is an expansion of precisely the form (1). If the left-hand member of (2) is not defined for a particular value $z = \beta_k$ it is to be defined as the limit of that left-hand member as $z$ approaches $\beta_k$. With this definition and under the hypothesis of Theorem I, the left-hand member of (2) is analytic even in the points $\beta_k$. It is a matter of indifference, as the reader will verify, whether $a_{n+k}$ is determined from (1) by the method previously described, or from (2) by that same method.

Theorem I is a special case of the following more general theorem, which we shall also prove.

THEOREM II. *If $f(z)$ is meromorphic for $|z| < T$, if the poles of $f(z)$ of modulus less than $T$ belong to the sequence $\alpha_1, \alpha_2, \ldots$, which has no limit point of modulus less than $A$, and where each pole occurs in the sequence at least a number of times corresponding to its multiplicity, and if finally the points $\beta_1, \beta_2, \ldots$ are distinct from the $\alpha_n$ and have no limit point of modulus greater than $B < A$, $T$, then there exists an expansion found by interpolation in the points $\beta_n$*

$$f(z) = a_0 + a_1\frac{z-\beta_1}{z-\alpha_1} + a_2\frac{(z-\beta_1)(z-\beta_2)}{(z-\alpha_1)(z-\alpha_2)} + \ldots, \qquad (3)$$

*valid for $|z| < (AT - BT - 2AB)/(A - B + 2T)$, $z$ distinct from the points $\alpha_n$. The series (3) converges absolutely for all values of $z$ (other than the $\alpha_n$) in modulus less than $(AT - BT - 2AB)/(A - B + 2T)$ and converges uniformly in any closed region interior to the circle $|z| = (AT - BT - 2AB)/(A - B + 2T)$ and containing no point $\alpha_n$.*

The special cases $A = \infty$ and $T = \infty$ are not excluded; Theorem I is the special case $A = \infty, T = \infty$.

Theorem II is to be proved by means of the following theorem.[1]

THEOREM III. *Let the function $f(z)$ be analytic for $|z| < T$ and let the numbers $\alpha_{in}(i = 1, 2, \ldots, n;\ n = 1, 2, \ldots)$ have no limit point whose modulus*

*is less than A, and the numbers* $\beta_{in} (i = 1, 2, \ldots, n + 1;\ n = 1, 2, \ldots)$ *no limit point whose modulus is greater than* $B < A$, $T$. *Then the sequence of rational functions (necessarily unique)*

$$f_n(z) = \frac{a_{0n}z^n + a_{n1}z^{n-1} + \ldots + a_{nn}}{(z - \alpha_{1n})(z - \alpha_{2n})\ldots(z - \alpha_{nn})} \tag{4}$$

*which coincide with the function* $f(z)$ *in the points* $\beta_{in}$, *approaches the limit* $f(z)$ *for* $|z| < (AT - BT - 2AB)/(A - B + 2T)$, *uniformly for* $|z| \leqq R < (AT - BT - 2AB)/(A - B + 2T)$.

If $f(z)$ is not *analytic* for $|z| < T$ but is *meromorphic*, with a finite number of poles of modulus less than $T$ distinct from the $\beta_{in}$, Theorem III may still yield a result on the expansion of $f(z)$, as we proceed to indicate. Let the first $m$ of the numbers $\alpha_{1n}\ \alpha_{2n}, \ldots, \alpha_{nn}$ be $\alpha_1, \alpha_2, \ldots, \alpha_m$, *independent of* $n$ $(n \geqq m)$, and let the poles of $f(z)$ of modulus less than $T$ be included in this set, each pole enumerated according to its multiplicity. Let the first $m'$ of the numbers $\beta_{1n}, \beta_{2n}, \ldots, \beta_{n+1,n}$ be $\beta_1, \beta_2, \ldots, \beta_{m'}$, also independent of $n (n \geqq m' - 1)$, $\beta_i \neq \alpha_j$. We assume the points $\alpha_{m+k,n} (k > 0)$ to have no limit point of modulus less than $A$ and the points $\beta_{m'+k,n} (k > 0)$ to have no limit point of modulus greater than $B$, and $f(z)$ is conveniently assumed analytic in all the points of interpolation $\beta_{in}$.

Let $R(z)$ be a rational function of $z$ of the form

$$R(z) = \frac{a_0z^{m'-1} + a_1z^{m'-2} + \ldots + a_{m'}}{(z - \alpha_1)(z - \alpha_2)\ldots(z - \alpha_m)},$$

which takes on the same value as $f(z)$ in the $m'$ points $\beta_1, \beta_2, \ldots, \beta_{m'}$; this function $R(z)$ is uniquely determined. The function

$$F(z) = \frac{(z - \alpha_1)(z - \alpha_2)\ldots(z - \alpha_m)}{(z - \beta_1)(z - \beta_2)\ldots(z - \beta_{m'})}\left[f(z) - R(z)\right]$$

is then analytic for $|z| < T$.

We approximate to $F(z)$ according to Theorem III by rational functions of the form

$$F_n(z) = \frac{b_{0n}z^{n-m'} + b_{1n}z^{n-m'-1} + \ldots + b_{n-m',n}}{(z - \alpha_{m+1,n})(z - \alpha_{m+2,n})\ldots(z - \alpha_{nn})}.$$

In the case $m \geqq m'$, this function $F_n(z)$ is precisely of form (4), where $m - m'$ of the (formal) poles are restricted to lie at infinity. In the case $m' > m$, this function $F_n(z)$ is of form (4) with the additional requirement

$$F_n(\infty) = 0,$$

this last condition given $m' - m$ times, equivalent to requiring a zero of $F(z)$ at infinity formally of order $m' - m$. The function $F_n(z)$ is required

to coincide with $F(z)$ in the points $\beta_{m'+1,n}\ \beta_{m'+2,n}, \ldots, \beta_{n+1,n}$, and is thereby uniquely determined. In the case $m \geqq m'$ it follows directly from Theorem III and in the case $m' > m$ can be proved in a manner similar to the proof of Theorem III (loc. cit.) that we have

$$\lim_{n \to \infty} F_n(z) = F(z)$$

uniformly for $|z| \leqq R < (AT - BT - 2AB)/(A - B + 2T)$. Then if we set

$$f_n(z) = R(z) + \frac{(z - \beta_1)(z - \beta_2) \ldots (z - \beta_{m'})}{(z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_m)} F_n(z) \tag{5}$$

$$= \frac{a_{0n}z^n + a_{1n}z^{n-1} + \ldots + a_{nn}}{(z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_m)(z - \alpha_{m+1,n}) \ldots (z - \alpha_{nn})}$$

we have

$$f(z) = \lim_{n \to \infty} f_n(z) \tag{6}$$

uniformly in any closed region interior to the circle $|z| = (AT - BT - 2AB)/(A - B + 2T)$ and containing no point $\alpha_1, \alpha_2, \ldots, \alpha_m$. The function $f_n(z)$ defined by (5) is of degree $n$, coincides with the function $f(z)$ in the points $\beta_1, \beta_2, \ldots, \beta_{m'}, \beta_{m'+1,n}, \beta_{m'+2,n}, \ldots, \beta_{n+1,n}$, and has its poles in the points $\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1,n}, \alpha_{m+2,n}, \ldots, \alpha_{nn}$. We have thus proved a theorem on the expansion of meromorphic functions, which the reader will readily state explicitly, and which includes Theorem III as a special case, dependent on the hypothesis that the first $m$ of the numbers $\alpha_{1n}, \alpha_{2n}, \ldots, \alpha_{nn}$ are independent of $n$ ($n \geqq m$) and that the first $m'$ of the numbers $\beta_{1n}, \beta_{2n}, \ldots, \beta_{n+1,n}$ are independent of $n (n \geqq m' - 1)$; on the other hand, *we do not here require, as in Theorem III, that the points $\alpha_{in}$ have no limit point of modulus less than A and that the points $\beta_{in}$ have no limit point of modulus greater than B so far as the $\alpha_n$ and $\beta_n$ themselves are concerned.*

In our discussion of development (6) we have for convenience supposed that the *first* $m$ of the numbers $\alpha_{1n}, \alpha_{2n}, \ldots, \alpha_{nn}$ are independent of $n (n \geqq m)$, but it is naturally sufficient if *any* $m$ of the numbers $\alpha_{1n}, \alpha_{2n}, \ldots, \alpha_{nn}$ are independent of $n$, for $n$ greater than some $M$. A corresponding remark holds with regard to the numbers $\beta_{in}$.

In either Theorem III or development (6), if the points $\alpha_{in}$ ($i \leqq n$) and $\beta_{in} (i \leqq n + 1)$ do not depend on $n$, the formal development of $f(z)$ is precisely of type (1); for the function $f_n(z) - f_{n-1}(z)$ is a rational function of $z$ of degree $n$ which vanishes in the points $z = \beta_1, \beta_2, \ldots, \beta_n$ and which has the (formal) poles $z = \alpha_1, \alpha_2, \ldots, \alpha_n$:

$$f_n(z) - f_{n-1}(z) = a_n \frac{(z - \beta_1)(z - \beta_2) \ldots (z - \beta_n)}{(z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_n)}.$$

Reciprocally, a formal development of type (1), where the coefficients are determined as we have described, is a formal development of type (4) where the equations

$$\beta_i = \beta_{in}, \qquad i \leqq n + 1,$$

$$\alpha_1 = \alpha_{1n}, \qquad i \leqq n,$$

are valid; the function $f_n(z)$ [sum of the first $n + 1$ terms of (1)] equals $f(z)$ in the points $\beta_{in}$.

The same convention that we have made for Theorem I is naturally assumed to hold for Theorems II and III and development (6) relative to interpolation in multiple points $\beta_{in}$.

Theorem II can now be proved either (*i*) by identifying (3) with (6) or (*ii*) by applying Theorem III directly as was done in the proof of (6). Let us indicate the details according to method (*ii*).

Choose arbitrarily $A' < A$, $T' < T$, $B' > B$, and choose $N$ so that for $n > N$ we have $|\alpha_n| > A', T'; |\beta_n| < B'$. All the coefficients $a_n$ in (3) are uniquely determined, and for $n > N$ the left-hand member of (2) is analytic for $|z| < T'$. Moreover, the points $\alpha_{in}$ (which may here be written $\alpha_n$, $n > N$) have no limit point of modulus less than $A'$, and the points $\beta_{in}$ (which may here be written $\beta_n$, $n > N$) have no limit point of modulus greater than $B'$. Then for a fixed $n$ greater than $N$ the equation (2) is valid, by Theorem III, uniformly for $|z| \leqq R < (A'T' - B'T' - 2A'B')/(A' - B' + 2T')$. It is now allowable to let $A', T', B'$ approach $A, T, B$, respectively; the validity of (2) and therefore of (3) follows, uniformly for $|z| \leqq R < (AT - BT - 2AB)/(A - B + 2T)$. For the absolute convergence of the series (2) and hence of (3) we refer to our detailed proof of Theorem III (loc. cit.).

In Theorem III, and hence in Theorem II and expansion (6), if we have $\beta_{1n} = 0$, $\beta_{in} = 1/\bar{\alpha}_{i-1,n}$, with $A, T > 1$, the limit $(AT - BT - 2AB)/(A - B + 2T)$ can everywhere be replaced by $(A^2T + T + 2A)/(2AT + A^2 + 1)$; if the given function $f(z)$ is analytic for $|z| \leqq 1$, the approximating functions are the respective functions of form (4) of best approximation to $f(z)$ on $|z| = 1$ in the sense of least squares, provided $|\alpha_{in}| > |$.

It may be noticed that in Theorem II and development (6) it is not necessary to assume $f(z)$ analytic in *all* the points of interpolation $\beta_n$. Indeed, so far as concerns the points $\beta_n$ of modulus not less than $T$, the condition $f_j(\beta_n) = f(\beta_n)$ has no necessary relation to $f(z)$ as a monogenic analytic function, and those conditions are essentially arbitrary auxiliary conditions. For instance, we may choose $k$ points $\beta_n$ at infinity and write $f_j(\infty) = 0$ as a $k$-fold condition, so that expansion (6) includes the extension of Theorem III used in studying the convergence of $F_n(z)$ for the case $m' > m$ in the proof of (6).

Series of form (1) have been previously studied by Angelescu,[2] but apparently only for analytic (not meromorphic) functions, and only under the hypothesis of the existence of $\lim_{n\to\infty} \alpha_n$ and $\lim_{n\to\infty} \beta_n$. The connection of such series with approximation in the sense of least squares, as used in Theorem I, seems first to have been pointed out by the present writer.[3]

An expansion of form (1) can readily represent the function zero in a region (not containing all of the points $\beta_n$) without having all of the coefficients $a_n$ to vanish, as Angelescu[2] has indicated. This follows easily likewise from our present discussion. For instance, let $f(z)$ be analytic for $|z| < T$, and let $\beta_1, \beta_2, \ldots, \beta_n$ be of modulus not less than $T$, while $\beta_{n+1}, \beta_{n+2}, \ldots$ are of modulus not greater than $B$. If the numbers $\alpha_1, \alpha_2, \ldots$, are all of modulus greater than $A$, the coefficients $a_0, a_1, \ldots, a_{n-1}$ in (2) may be chosen *arbitrarily;* the expansion (2) is valid, as follows from Theorem III, and leads to an expansion (1) for $f(z)$ also valid uniformly for $|z| \leqq R < (AT - BT - 2AB)/(A - B + 2T)$. In particular, the expansion is valid if $f(z)$ vanishes identically; the coefficients $a_0, a_1, \ldots, a_{n-1}$ are arbitrary.

By an arbitrary linear substitution

$$w = \frac{\alpha z + \beta}{\gamma z + \delta}, \qquad \alpha\delta - \beta\gamma \neq 0,$$

the expansion (3) is transformed into an expansion of precisely the same form. In general, the convergence to $f(z)$ at $z$ of an expansion (3) as used in Theorem III depends on the analytic or meromorphic character of $f(z)$ in all points $t$ such that we have

$$|(t, \alpha, z, \beta)| = \left|\frac{(t - \alpha)(z - \beta)}{(z - \alpha)(t - \beta)}\right| \geqq 1,$$

where $\alpha$ and $\beta$ represent respectively arbitrary limit points of the sets $\alpha_n$ and $\beta_n$. Corresponding to this fact, a more general theorem than Theorem III can be expressed, which, moreover, is in a form invariant under linear transformation; compare the reference already given.

The restriction $\alpha_i \neq \beta_j$, which we have constantly made, is a matter of convenience rather than necessity. Many of our results are still valid if that restriction is removed.

[1] Walsh, *Trans. Amer. Math. Soc.*, **34** (1925).

[2] Angelescu, *Bull. Acad. Roumaine*, **9**, 164–168 (1925).

[3] Compare, however, Malmquist, *Compt. rend. du sixième congrès (1925) des mathématiciens scandinaves*, Copenhagen, 1926, pp. 253–259.

# CONTRIBUTION TO THE MATHEMATICAL THEORY OF CAPTURE. I. CONDITIONS FOR CAPTURE

BY A. J. LOTKA

NEW YORK CITY

Communicated January 12, 1932

The equations for a simple biological system comprising one *prey* species and one *predatory* species, the latter consuming the former for food, have been given by the writer (1920, 1923, 1925), and, independently, by Volterra (1926) in the form

$$\frac{dX_1}{dt} = aX_1 - kX_1X_2$$
$$\frac{dX_2}{dt} = kk'X_1X_2 - bX_2, \tag{1}$$

where $X_1$, $X_2$ are, respectively, the total masses of the two species $S_1$ and $S_2$; or, with slight and obvious change in the argument and in the physical significance of the coefficients, $X_1$ and $X_2$ may be taken to represent the number of individuals in the two species. We shall here read them in this latter sense. The coefficients $a$, $k$, $k'$, $b$, may, in first approximation, and especially for small ranges of variation of $X_1$ and $X_2$, be regarded as constants; or, in more exact treatment of the problem, these coefficients are themselves functions of $X_1$ and $X_2$ (Lotka, 1923; Volterra, 1926).

The purpose of the present communication is to single out for further discussion the term $kk'X_1X_2$ in (1), and, in particular, the coefficient $k$. The physical significance of this coefficient depends on the particular physical process by which individuals of the species $S_2$ "consume" those of species $S_1$. We shall here consider the case in which this takes place by pursuit and capture, an individual $S_1$ being killed, at capture, by an individual $S_2$, and consumed in whole or in part. The further resolution of the coefficient $k$ is the analytical counterpart of the further resolution of the physical process of capture.

*Influence of Topography: Distributed Refuges.*—Now capture can take place in various ways, as, for example, by simple pursuit until the pursuer overtakes the pursued. In such case capture *can* take place only if the velocity of the pursuer exceeds that of the pursued; and *must* take place, in that event, provided sufficient time is given. But this is not the typical situation in nature. What usually happens is that the pursued runs to *cover*, i.e., to a *refuge*, and escapes if it reaches effective cover before being overtaken.

*Cover* can be of various kinds. In the case of a bird, for example, rising from the ground affords effective refuge from terrestrial enemies. But a case of particular interest is that in which cover is presented by features of the topography. The problem of the frequency of capture in this case resolves itself into a triangular study of the relations between three classes of factors, namely,

1. Properties of the individuals $S_1$.
2. Properties of the individuals $S_2$.
3. Properties of the topography over which the two species operate.

Preserving the essential characteristics of typical organisms, but reducing them to their simplest terms for the purpose of achieving a first approach to the problem, we shall conceive the process of capture as follows:

1. *Random Motion between Encounters.*—Every individual $S_1$ and $S_2$ moves "in random manner" over the topography so long as it is not "in encounter" with an individual of the other species. As a first approximation we shall assume that the random motion is rectilineal, the azimuth changing at each encounter, all azimuths being equally frequent, and with no correlation between azimuths before and after encounter.

2. *Encounters, First Phase: The "Stalk."*—About each individual $S_2$ a field can be described, which we shall speak of as the *field* of *influence* of $S_2$. In general this field might have various shapes, and might be a function of the topography; but still restricting ourselves to the simplest case, we shall suppose that the field is a circle of radius $r_2$ and center attached to $S_2$. This field of influence has the following property: As soon as an individual $S_1$ enters this field $r_2$, the motion of $S_2$ is, *in a fraction $\psi$ of such cases*, no longer random, but follows a "curve of pursuit" defined as a characteristic property of the species $S_2$. The fraction $\psi$ measures the *observability* (*visibility*) of $S_1$ with respect to $S_2$. This observability depends, in a manner which we shall not here seek to analyze further, (*a*) on features of the environment; and (*b*) on features of the prey species (coloration, mimicry, camouflage, etc.).

3. *Encounter, Second Phase: The Pursuit or Flight.*—The motion of $S_2$ continues as stated under (2) above, until in turn $S_2$ enters the field of influence of $S_1$. For simplicity we shall here assume that this also is circular, with radius $r_1$ and center at $S_1$. From the moment that $S_2$ enters into the field of $S_1$, e.g., as soon as the prey senses (sights) the pursuer, the motion of $S_1$, which until then may have had any random character, follows, in a fraction $\chi$ of such cases, *a curve of flight* characteristic of the species $S_1$. The fraction $\chi$ measures the observability (visibility) of $S_2$ with respect to $S_1$.

*Distribution of Cover or Refuges.*—In the most general case the topography of the system may be complex, and may require, for its complete

definition, a map indicating the character of the terrain at every point, as a function of the coördinates of position. For the present, however, we shall restrict ourselves to features of the topography which can be defined in statistical terms, and, in particular, we shall consider the influence of *cover* or *refuge* upon the course of events. In general refuges may be distributed in any manner over the topography, with a density $\delta$ per unit area, and the entire field may be divided into separate domains, one domain for each refuge, such that every point within a domain is nearer to the refuge (considered as a point) of that domain, than to any other refuge. So, for example, in figure 1, the points $P_1$, $P_2$. . . represent refuges. Each is contained within a polygon whose sides are formed by the perpendicular bisectors of the joins of adjacent refuge points. For example, the point $Q_1$ is nearer to $P_1$ than to any other refuge.

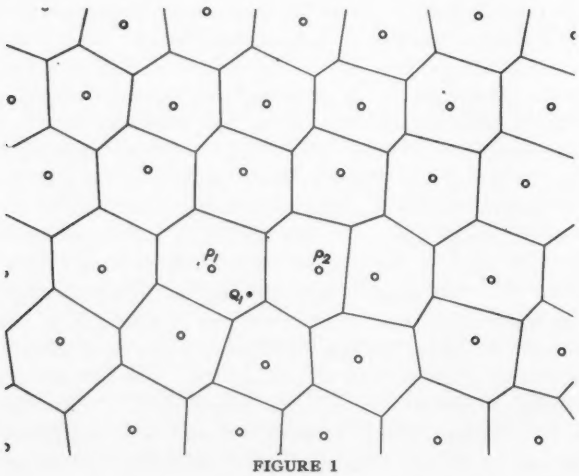



FIGURE 1

The perfectly general case of distribution of refuges may be narrowed down in various ways; for example, the refuge density, though variable over small areas, may be (in the limit) uniform, provided a sufficiently large area is observed. For a restricted patch, of area $A$, the frequency $f(\Delta)$ of the deviation $\Delta$ of the actual density from the mean taken over a large field, may be defined by some such law as

$$f(\Delta) = ke^{-\varphi(A)\,\Delta^2} \qquad (2)$$

where $\varphi(A)$ becomes very large when $A$ is not small, so that for larger areas there are practically no deviations of the average density, from

the mean over the entire terrain. This suggestion is here thrown out by the way of a hint for further generalization.

For the present we shall restrict our discussion to cases in which the *distribution pattern* does not materially affect our results, or to conventionalized cases in which the distribution pattern is simple, as will be indicated more in detail later. But first we must isolate for examination a single element of the general process under discussion. It is to this that the following section is devoted.

*Single Refuge, Single Encounter: Conditions for Capture.*—It will be convenient, in the following development, to consider first the simpler case in which $\psi = \chi = 1$, that is, in which every individual (prey or predator) that comes within the radius of observation of its opponent, is actually observed by that opponent (predator or prey).

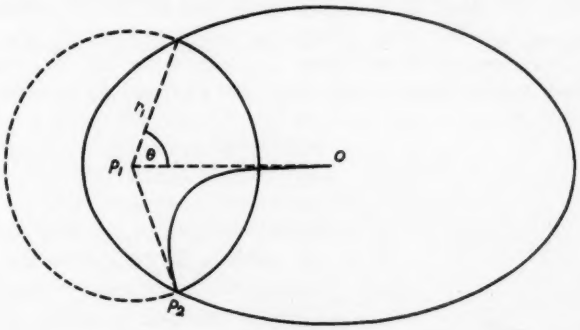



FIGURE 2

Consider first, then, a single refuge located at the point $O$ in figure 2. Let $P_1$, $P_2$, respectively, be the positions of $S_1$, $S_2$ at the beginning of the *second* phase (flight) of an encounter between them. Then $S_1$ immediately flees toward $O$ along $P_1O$, with velocity $v_1$, while $S_2$ pursues it with a velocity $v_2$ directed always toward $S_1$, that is, toward the moving point $P_1$. The point $P_2$ then traces a curve of the kind known as a *curve of pursuit.* In figure 3 a family of such curves of pursuit, all terminating in the point $O$, are shown. These correspond, then, to paths traced by $S_2$ in the limiting case that capture occurs just as $S_1$ reaches the refuge.

Now it can be shown* that the *isochrone* through $P_1$, that is, the locus of points corresponding to equal times upon the several curves of pursuit terminating synchronously at $O$, is an ellipse of eccentricity $e = \frac{v_1}{v_2}$ with one focus at $P_1$ and its center at $O$. (See figures 2 and 3.) For this ellipse the polar equation, dated from $P_2$ as pole and $P_1O$ as axis, is

$$\rho = \frac{Re}{1 - e^2}(1 - e\cos\theta) \tag{3}$$

or

$$R = \frac{\rho(1 - e^2)}{e(1 - e\cos\theta)} \tag{4}$$

where $\rho$ is the focal distance $P_1O$, while $\theta$ is the phase or amplitude $P_2P_1O$ in figure 2, and $R$ is the radius vector $P_2P_1$.

But, at the beginning of the second phase (flight) of the encounter, $P_2$ is at a distance $r_1$ (the range of perception of $S_1$, the pursued species) from $P_1$. In other words, at that instant $P_2$ lies somewhere on a circle drawn about $P_1$ as center, with radius $r_1$. If it lies at the intersection of this circle with the ellipse (3), capture will just take place as $S_1$ reaches the refuge at $O$. If $P_2$, on the other hand, lies anywhere in the arc of the circle drawn in a solid line in figure 3, the part within the ellipse, then, evidently, capture will occur before the refuge is reached; and on the contrary, if $P_2$ lies in the arc shown in a dashed line, the arc outside the ellipse, then $S_1$ will escape to the refuge, and there will be no capture.

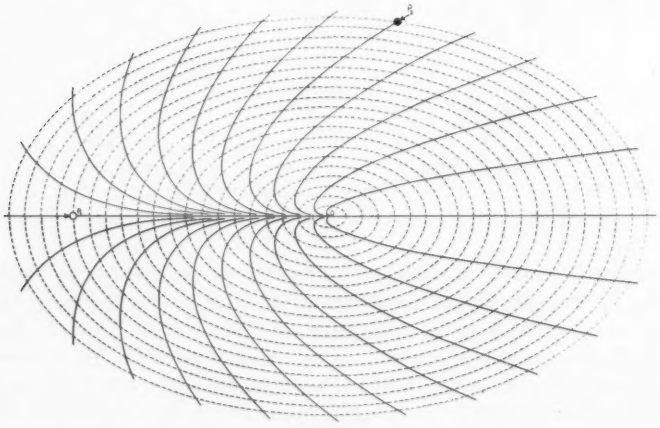


FIGURE 3

To sum up, capture will, or will not, take place, according as, at the moment pursuit begins, the pursuer lies within or lies outside an ellipse (3) of eccentricity $e = \frac{v_1}{v_2}$, drawn with the refuge $O$ as center, and the initial position $P_1$ of $S_1$ as focus. Analytically this statement takes the form that capture will, or will not, take place, according as

$$r_1 \lesseqgtr R \tag{5}$$

i.e., as

$$r_1 \lesseqgtr \frac{\rho(1 - e^2)}{e(1 - e\cos\theta)} \tag{6}$$

or, again, according as

$$\cos\theta \gtreqless \frac{1}{e} - \frac{\rho(1 - e^2)}{r_1 e^2} \tag{7}$$

or, as

$$\theta \lesseqgtr \text{arc cos}\left\{\frac{1}{e} - \frac{\rho(1 - e^2)}{r_1 e^2}\right\}. \tag{8}$$

It will be well to state the meaning of this relation in words. If at the beginning of the pursuit the pursued $S_1$ is at a distance $\rho$ from the refuge, then *capture* will occur if the angle at which $S_1$ sights its pursuer $S_2$ is *less than* the critical angle arc cos$\left\{\frac{1}{e} - \rho\frac{(1 - e^2)}{r_1 e^2}\right\}$, this angle being measured from the straight line $P_1O$ drawn to the refuge, from $S_1$, as indicated in figure 3. Capture will *not* take place, but, on the contrary $S_1$ will escape, if the angle mentioned is *greater* than the critical angle of capture. Evidently, as $\rho$, the distance from the refuge, increases, $\theta$, the critical angle within which capture is possible, increases also. These facts are brought out in the series of diagrams (Fig. 4).

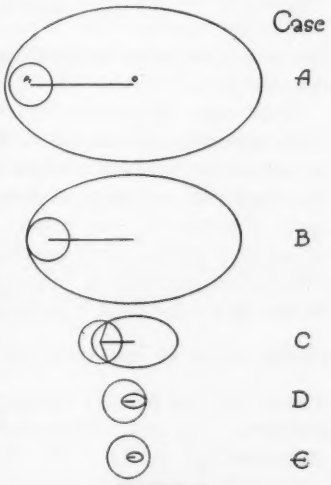


FIGURE 4

*Five Types of Encounters, and Their Analytical Characteristics: Safe, Dangerous and Fatal Encounters; also Two Transition Types.*—As has already been remarked, it is clear from figure 3 that of all possible azimuths of the position $P_2$ of $S_2$ with respect to $P_1O$ at the beginning of the pursuit, those represented by the dashed portion of the circle result in escape, those represented by the fully drawn portion result in capture.

A number of different cases will present themselves, as follows:

*Division of Entire Territory into Safe Field, Danger Zone and Fatal Field.*

1. If $\rho < \frac{r_1 e}{1 + e}$ the ellipse lies wholly within the circle, without touching

it. In that case capture is impossible, escape is certain; $S_1$ is *safe*. (Fig. 4*E*.)

2. If $\rho > \frac{r_1 e}{1 - e}$, the circle lies wholly within the ellipse, and capture is certain; $S_1$ is in a *fatal* position. (Fig. 4*A*.)

3. If $\frac{r_1 e}{1 + e} < \rho < \frac{r_1 e}{1 - e}$, then the ellipse and the circle intersect. This is the case already mentioned, in which capture takes place for some azimuths, but not for others, namely, capture takes place whenever

$$\theta > \arccos\left\{\frac{1}{e} - \frac{\rho}{r_1}\,\frac{1 - e^2}{e^2}\right\}$$

Here $S_1$ is in *danger* of capture. (Fig. 4*C*.)

4. Limiting cases arise if

(*a*) $\rho = \frac{r_1 e}{1 + e}$; the ellipse is then internally tangent to the circle. Capture is possible (for azimuth exactly $\theta = 0$), but is infinitely improbable. (Fig. 4*D*.)

(*b*) $\rho = \frac{r_1 e}{1 - e}$; the circle is then internally tangent to the ellipse. Escape is possible (for azimuth exactly $\theta = \pi$), but is infinitely improbable. (Fig. 4*B*.)

These cases taken together evidently form a graded series, of which a few members are shown in figure 4 to illustrate these relations. If we consider a series of positions of $P_1$ successively nearer to the refuge, the ellipse, at first large and enclosing the circle, gradually shrinks, till there is contact at one point. On its further shrinking there is, first intersection at two closely neighboring points, so that the angle $\theta$ is small; with further approach of $P_1$ to the refuge the points of intersection spread farther apart, the angle $\theta$ increases, and with it the probability of escape. Finally, when the distance of $P_1$ from the refuge has fallen to $\frac{r_1 e}{1 + e}$, there is again only one point of intersection (contact), the angle $\theta$ is $\pi$, and escape is assured. On still further approach the ellipse shrinks into the interior of the circle.

* A. J. Lotka, *Am. Math. Month.*, p. 421, 1928. Quite recently the same result has been published independently by V. Lalan, *Compt. Rend.*, **192**, 468, 1931.

# A NEW SET OF INDEPENDENT POSTULATES FOR THE ALGEBRA OF LOGIC WITH SPECIAL REFERENCE TO WHITEHEAD AND RUSSELL'S PRINCIPIA MATHEMATICA.*

BY EDWARD V. HUNTINGTON

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated January 14, 1932

The following set of independent postulates for Boolean algebra is expressed in terms of "logical addition" and "logical negation," and is of particular interest in connection with B. A. Bernstein's recent paper on "Whitehead and Russell's Theory of Deduction as a Mathematical Science" (*Bulletin* of the American Mathematical Society, June, 1931).

All the primitive propositions of Whitehead and Russell's Principia Mathematica, when expressed in Bernstein's "mathematicized" form, are deducible from the present set of postulates, so that the present set of postulates may be regarded as one of the simplest and most natural sets "from which all mathematics may be deduced."

The "universe of discourse" is the class of all systems $(K, +, ')$, where $K$ = a class of undefined elements, $a, b, c, \ldots$ (interpretable as regions in a square), while $a + b$ = the result of an undefined binary operation (interpretable as the smallest area which includes both $a$ and $b$), and $a'$ = the result of an undefined unary operation (interpretable as the area complementary to $a$ with respect to the square).

The postulates are the following six. (A trivial postulate requiring the class $K$ to contain at least two elements is assumed without further mention; and in Postulates 3–6 it is assumed that the indicated combinations are elements of $K$.)

*Postulate 1.* *If $a$ and $b$ are elements of $K$, then $a + b$ is an element of $K$.*
*Postulate 2.* *If $a$ is an element of $K$, then $a'$ is an element of $K$.*
*Postulate 3.* $a + b = b + a$.
*Postulate 4.* $(a + b) + c = a + (b + c)$.
*Postulate 5.* $a + a = a$.
*Postulate 6.* $(a' + b')' + (a' + b)' = a$.

If the "logical product," $ab$, is defined by $ab = (a' + b')'$, and the relation of "subsumption," $a < b$, is defined by writing $a < b$ when and only when $a + b = b$, then all the properties of Boolean algebra can be deduced from these six postulates. In particular, there is a unique element $z$ such that $a + z = a$, and a unique element $u$ $(= z')$ such that $a + a' = u$; and if the number of elements in $K$ is finite, it must be a power of 2.

The examples used to establish independence, together with the de-

tailed proof that this set of postulates is equivalent to the sets given by the present writer in 1904 (and to the sets given by H. M. Sheffer in 1913 and by B. A. Bernstein in 1916), will be included in a longer paper of which the present note is an abstract.

Two points may be added here.

In the first place, it can be shown that the Principia-Bernstein list is not sufficient to define Boolean algebra without the addition of a further postulate, which may be stated in various forms. The following form was suggested to me by Mr. P. Henle:

1.73. If $a' + b = 1$ *and* $b' + a = 1$, *then* $a = b$. (Here "1" is used to denote the "universe-element" of the system.)

In the second place, while the authors of the Principia base their theory "officially" on the operations of addition $(+)$ and negation $(')$, they introduce at the start, and use continually, another operation, $a \supset b$, defined by the equation $a \supset b = a' + b$. It may be of interest, therefore, to note that it is possible to construct a set of postulates for Boolean algebra which is expressed wholly in terms of $(K, \supset)$, that is, wholly in terms of what Whitehead and Russell call "implication."

Such a set is the following, in which $K$ is to be thought of as an undefined class of elements, $a, b, c, \ldots$, and $\supset$ as an undefined binary operator ("horse-shoe").

1′. *If* $a$ *and* $b$ *are elements of* $K$, *then* $a \supset b$ *is an element of* $K$.

2′. *There exists a unique element* $z$ *such that* $z \supset a = a \supset a$ *for every element* $a$.

*Definition.* $a \supset b \supset c = (a \supset b) \supset c$.

3′ $a \supset b \supset b = b \supset a \supset a$

4′. $a \supset b \supset b \supset c \supset c = a \supset (b \supset c \supset c) \supset (b \supset c \supset c)$.

5′. $a \supset a \supset a = a$.

6′. $a \supset (b \supset z) \supset z \supset (a \supset b \supset z) \supset (a \supset b \supset z) = a$.

7′. $a \supset z \supset b \supset b = a \supset b$.

If $a + b$ and $a'$ are defined in terms of $\supset$ by the equations $a + b = (a \supset b) \supset b$ and $a' = a \supset z$, it can be shown that this peculiar set of postulates is equivalent to the earlier sets. But the much greater complexity of Postulates 1′–7′ as compared with the "naturalness" of Postulates 1–6 may be taken as an indication that the choice of $(K, \supset)$ as the basis of the algebra would be far more "artificial" than the choice of $(K, +, ')$.

* Presented to the American Mathematical Society, December, 1931, and since revised.

# PRELIMINARY NOTE ON THE INVERSION OF THE LAPLACE INTEGRAL

BY D. V. WIDDER

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated January 15, 1932

In a previous note in these PROCEEDINGS[1] we stated the following result: *If* $\varphi(t)$ *is continuous in the interval* $0 \leqq t < \infty$, *if* $\lim_{t=\infty} \varphi(t) = a$, *and if*

$$f(x) = \int_0^\infty e^{-xt}\varphi(t)dt,$$

*then*

$$\lim_{k=\infty}\left[x^{k+1}\frac{(-1)^k f^{(k)}(x)}{k!} - \varphi\left(\frac{k}{x}\right)\right] = 0 \qquad (1)$$

*uniformly in the interval* $0 \leqq x < \infty$.

This result was obtained in order to discuss the zeros of $\varphi(t)$ in terms of those of the derivatives of $f(x)$. The result has, however, an important interest in itself, since it enables us to invert the Laplace integral (1) under the conditions described in the theorem. We have, in fact,

$$\varphi(t) = \lim_{k=\infty}(-1)^k \frac{f^{(k)}(k/t)k^{k+1}}{k!\,t^{k+1}}$$

uniformly for $0 \leqq t < \infty$.

This result was obtained earlier by E. Post[2] under the restriction of continuity on $\varphi(t)$ but without the condition that $\varphi(t)$ approaches a limit as $t$ becomes infinite. On the other hand the uniformity of the approach was not obtained by Post. In the present note we sketch a theory whereby the integral equation (1) may be solved with *no* restriction on $\varphi(t)$ as to continuity. We simply impose the natural condition that $\varphi(t)$ shall be integrable in the sense of Lebesgue and (so that the integral (1) may converge for $x$ sufficiently large) the condition

$$|\varphi(t)| < Me^{\sigma t} \quad 0 < t < \infty,$$

where $M$ and $\sigma$ are constants.

We then turn to the problem of inverting integrals of the form

$$f(x) = \int_0^\infty e^{-xt}d\alpha(t), \qquad (2)$$

where *no* restrictions are imposed on $\alpha(t)$ except those necessary to make the improper Stieltjes integral have a meaning: $\alpha(t)$ is of bounded varia-

tion in every finite interval and is such that (2) converges for some $x$ sufficiently large. It is important to note that the absolute convergence of (2) is not required. The result obtained is the following:

$$\frac{\alpha(t+) + \alpha(t-)}{2} = \lim_{k=\infty}\left[ f(\infty) + \int_{k/t}^{\infty} \frac{(-1)^{k+1} u^k f^{(k+1)}(u)}{k!}\, du \right]. \quad (3)$$

These results form a most powerful tool for the solution of the type of problem considered by the author in an earlier paper.[3] For example, consider Bernstein's theorem proved there. It states that if $f(x)$ has its derivatives alternately positive and negative in the interval $0 < x < \infty$, then $f(x)$ can be represented in the form (2) with $\alpha(t)$ a non-decreasing function. This result could be predicted at once from (3). For under the assumed conditions on $f(x)$ the integrand in (3) is positive so that the quantity in brackets is quite obviously an increasing function of $t$. Hence the limit $\alpha(t)$ will be non-decreasing if it exists. We are able to show further that the limit does exist and thus give an elementary proof of Bernstein's theorem involving no use of the moment problem. Further applications of the result are made to the determination of necessary and sufficient conditions for the representation of $f(x)$ in the form (2) where $\alpha(t)$ is of prescribed form. The following theorems serve to sketch the theory.

THEOREM 1. *The limit*

$$\lim_{k=\infty} e^{-k/t}\left[ 1 + \frac{k}{t} + \frac{1}{2!}\left(\frac{k}{t}\right)^2 + \ldots + \frac{1}{k!}\left(\frac{k}{t}\right)^k \right]$$

*exists and has the value*

$$\begin{array}{ll} 0 & 0 < t < 1, \\ 1/2 & t = 1, \\ 1 & t > 1. \end{array}$$

This theorem is proved by obtaining an integral expression for the quantity in brackets and then applying known methods to the determination of an asymptotic expression for it.

THEOREM 2. *If*

$$f(x) = \int_0^{\infty} e^{-xt} \varphi(t) dt,$$

*where the integral converges for $x$ sufficiently large and $\varphi(t)$ is of bounded variation in every finite interval, then*

$$\lim_{k=\infty} \frac{(-1)^k f^{(k)}(k/t) k^{k+1}}{k!\, t^{k+1}} = \frac{\varphi(t+) + \varphi(t-)}{2}.$$

By use of this theorem we are able to prove

THEOREM 3. *If*

$$f(x) = \int_0^\infty e^{-xt}\, d\alpha(t),$$

*where $\alpha(t)$ is of bounded variation in every finite interval, the integral converging for $x$ sufficiently large, then*

$$\lim_{k=\infty}\left[f(\infty) + \int_{k/t}^\infty \frac{(-1)^{k+1}u^k f^{k+1}(u)}{k!}\, du\right] = \frac{\alpha(t+) + \alpha(t-)}{2}.$$

It is possible to obtain another expression for $\alpha(t)$ depending on the values of $f^{(k)}(x + c)$, where $c$ is any constant. It is

$$\lim_{k=\infty}\left[f(\infty) + (-1)^{k+1}\int_{k/t}^\infty \frac{u^k}{k!}\, e^{ck/u} f^{(k+1)}(u + c)du\right] = \frac{\alpha(t+) + \alpha(t-)}{2}.$$

This clearly reduces to the result of Theorem 3 when $c$ is zero.

THEOREM 4. *Let*

$$f(x) = \int_0^\infty e^{-xt}\varphi(t)dt,$$

*where $\varphi(t)$ is integrable in the sense of Lebesgue and uniformly bounded in the interval $(0, \infty)$. Then*

$$\lim_{k=\infty} \frac{(-1)^k f^{(k)}(k/t)k^{k+1}}{k!\, t^{k+1}} = \varphi(t)$$

*almost everywhere.*

This theorem can easily be extended to the case in which $\varphi(t)$ is not uniformly bounded but satisfies a condition of the type described in the introduction. We give next two uniqueness theorems.

THEOREM 5. *Let $f(x)$ satisfy the condition*

$$|f^{(k)}(x)| < \frac{Mk!}{x^{k+1}},\ x > 0;\ k = 0, 1, 2, \ldots$$

*Then*

$$\lim_{k=\infty} (-1)^k \int_0^\infty \frac{e^{-xt} f^{(k)}(k/t)k^{k+1}}{k!\, t^{k+1}}\, dt = f(x),\ x > 0.$$

THEOREM 6. *Let $f(x)$ satisfy the condition*

$$\left|\int_x^\infty \frac{u^k}{k!} f^{(k+1)}(u)du\right| < M, \quad x > 0;\ k = 0, 1, 2, \ldots,$$

*and let*

$$\varphi_k(t) = (-1)^{k+1} \int_{k/t}^{\infty} \frac{u^k}{k!} f^{(k+1)}(u)du + f(\infty).$$

*Then,* if $\varphi_k(t)$ is of bounded variation in every finite interval,

$$\lim_{k=\infty} \int_0^{\infty} e^{-xt} d\varphi_k(t) = f(x), \quad x > 0.$$

Both of these theorems can be made more general so as to obtain an expression for $f(x)$ in an arbitrary interval.

By use of these theorems we can prove the following results.

THEOREM 7. *A necessary and sufficient condition that $f(x)$ should be completely monotonic for $x > c$ is that*

$$f(x) = \int_0^{\infty} e^{-xt} d\alpha(t),$$

*where $\alpha(t)$ is non-decreasing and the integral converges for $x > c$.*

THEOREM 8. *A necessary and sufficient condition that $f(x)$ can be represented in the form*

$$f(x) = \int_0^{\infty} e^{-xt} d\alpha(t),$$

*the integral converging absolutely for $x > 0$, is that*

$$\int_x^{\infty} \frac{u^k}{k!} \mid f^{(k+1)}(u) \mid du < M \quad x > 0; \quad k = 0, 1, 2, \ldots.$$

For earlier proofs of Theorem 7 and Theorem 8 reference may be made to the author's paper already cited.[3] The proofs now obtained have distinct advantages over those given earlier in that they involve no reference to the moment problem.

[1] D. V. Widder, "On the Changes of Sign of the Derivatives of a Function Defined by a Laplace Integral," *Proc. Nat. Acad. Sci.*, **18**, 112–114(1932).

[2] E. L. Post, "Generalized Differentiation," *Trans. Am. Math. Soc.*, **32**, 723 (1930).

[3] D. V. Widder, "Necessary and Sufficient Conditions for the Representation of a Function as a Laplace Integral," *Trans. Am. Math. Soc.*, **33**, 851 (1931).

# ON THE UNITARY INVARIANTS OF A SQUARE MATRIX

BY FRANCIS D. MURNAGHAN

DEPARTMENT OF MATHEMATICS, THE JOHNS HOPKINS UNIVERSITY

Communicated January 6, 1932

Let $\mathbf{A}$ be a square $n$-rowed matrix with complex (including real) elements and let $(\lambda, \ldots, \lambda_n)$ denote its characteristic numbers. These characteristic numbers furnish the complete system of invariants of $\mathbf{A}$ under transformation of $\mathbf{A}$ by *any* non-singular matrix but there are in general other invariants when the transforming matrix is restricted to be unitary and it is the purpose of this note to give the complete system of unitary invariants. In the general (= non-special) case there are $(n-1)^2$ real unitary invariants in addition to the $2n$ real invariants furnished by the spectrum of characteristic numbers. In the course of the argument will be given an explicit evaluation of Schur's triangular canonical form of an arbitrary square matrix under unitary transformations, and the close connection between this normalization and the classical orthogonalization process of Schmidt will be made evident.

To facilitate the exposition we shall first consider the non-special case where all the characteristic numbers are different. To each characteristic number $\lambda_p$ will correspond a characteristic vector $\mathbf{x}_p$ which may be supposed normalized to the extent of having its magnitude unity, i.e., $(\mathbf{x}_p | \mathbf{x}_p) \equiv x_p^1 \bar{x}_p^1 + \ldots + x_p^n \bar{x}_p^n = 1$, but there will remain an arbitrary scalar factor $\tau_p$ of modulus unity on each characteristic vector $\mathbf{x}_p$. If we denote by $\mathbf{X}$ the characteristic matrix (whose columns are the characteristic vectors) $\mathbf{X}$ is non-singular (since the characteristic vectors are linearly independent). $\mathbf{X}$ is indeterminate to the extent of a post-factor phase matrix $\mathbf{F}$ (a diagonal unitary matrix) whose diagonal elements are the indeterminate scalar factors $(\tau_1, \tau_2, \ldots, \tau_n)$ just referred to. In other words if $\mathbf{X}$ is a characteristic matrix the general characteristic matrix is $\mathbf{X}_1 = \mathbf{XF}$. The second norm $\mathbf{N} = \mathbf{X}^*\mathbf{X}$ (where $\mathbf{X}^*$ is the conjugate transposed matrix to $\mathbf{X}$) is indeterminate to the extent $\mathbf{N}_1 = \mathbf{F}^*\mathbf{NF}$ (the first norm $\mathbf{XX}^*$ being uniquely determined). The first result to be established is that two non-special square matrices $\mathbf{A}$ and $\mathbf{B}$ are unitarily equivalent when and only when they have the same spectrum $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and when the second norms $\mathbf{X}^*\mathbf{X}$ and $\mathbf{Y}^*\mathbf{Y}$ of any pair of their characteristic matrices (the order in which the characteristic vectors are chosen being determined by the ordering of the characteristic numbers) are connected by the relation

$$\mathbf{Y}^*\mathbf{Y} = \mathbf{F}^*\mathbf{X}^*\mathbf{XF}; \ (\mathbf{F} \text{ an arbitrary phase matrix}). \qquad (1)$$

To demonstrate this result we first recall that the positively definite

matrix $\mathbf{X^*X}$ may be factored by the classical method (Jacobi-Toeplitz)[1] into the form $\mathbf{T^*T}$ where

$$\mathbf{T} = \begin{pmatrix} t_1^1 & t_2^1, & \dots & t_n^1 \\ 0 & t_2^2, & \dots, & t_n^2 \\ 0, & & \dots, & \\ 0, & 0, & \dots, & t_n^n \end{pmatrix}.$$

is a triangular matrix. To see the degree of arbitrariness in this factorization let $\mathbf{X^*X} = \mathbf{T_1^*T_1} = \mathbf{T_2^*T_2}$ be two such factorizations; introducing the polar representation described in a previous paper[2] we have $\mathbf{T_1} = \mathbf{U_1P}$, $\mathbf{T_2} = \mathbf{U_2P}$, where $\mathbf{U_1}$ and $\mathbf{U_2}$ are unitary while $\mathbf{P}$ is the uniquely determinate positively definite matrix whose square $= \mathbf{X^*X}$. Eliminating $\mathbf{P}$ we have $\mathbf{T_2T_1^{-1}} = \mathbf{U_2U_1^*}$ and this product (being both triangular and unitary) must be a diagonal phase matrix $\mathbf{D}$. The indeterminate factor $\mathbf{D}$ in the equation $\mathbf{T_2} = \mathbf{DT_1}$ may be used to normalize uniquely $\mathbf{T_2}$ so that it has its diagonal elements real and positive. Hence there is a *uniquely determinate* triangle matrix with real and positive diagonal elements such that

$$\mathbf{X^*X} = \mathbf{T^*T}.$$

The polar representations $\mathbf{X} = \mathbf{U_1P}$; $\mathbf{T} = \mathbf{U_2P}$ show that

$$\mathbf{X} = \mathbf{UT}. \tag{2}$$

This equation gives an explicit formulation of Schmidt's orthogonalization process (normalized by the hypothesis that the diagonal elements of $\mathbf{T}$ are real and positive). Now the definition of characteristic vectors is equivalent to the equation $\mathbf{AX} = \mathbf{XL}$, where $\mathbf{L}$ is the diagonal matrix whose elements are the characteristic numbers $(\lambda_1, \dots, \lambda_n)$ of $\mathbf{A}$. On using (2) we obtain at once

$$\mathbf{U^*AU} = \mathbf{TLT^{-1}}. \tag{3}$$

The matrix on the right of this equation is a triangle matrix, being the product of three triangle matrices, so that (3) furnishes explicitly Schur's reduction of $\mathbf{A}$ to the canonical triangular form by means of a unitary matrix. It is easy now to prove the theorem previously stated. As for the necessity we know that if $\mathbf{A}$ and $\mathbf{B}$ are unitarily equivalent: $\mathbf{B} = \mathbf{UAU^*}$, a characteristic matrix $\mathbf{Y}$ of $\mathbf{B}$ is furnished by the equation $\mathbf{Y} = \mathbf{UX}$ which implies immediately $\mathbf{Y^*Y} = \mathbf{X^*X}$. The sufficiency is proved as follows: from $\mathbf{Y^*Y} = \mathbf{F^*X^*X\,F}$ and the uniqueness of the normalized $\mathbf{T}$ it follows that $\mathbf{Y} = \mathbf{VT_1}$, $\mathbf{X\,F} = \mathbf{WT_1}$ ($\mathbf{V}$, $\mathbf{W}$ unitary). Combining these with $\mathbf{X} = \mathbf{UT}$ we have

$$\mathbf{T}_1 = \mathbf{U}_1\mathbf{T}\mathbf{F};\ \mathbf{U}_1 = \mathbf{W}^*\mathbf{U}.$$

Hence $\mathbf{T}_1\mathbf{L}\mathbf{T}_1^{-1} = \mathbf{U}_1\mathbf{T}\mathbf{L}\mathbf{T}^{-1}\mathbf{U}_1^*$ ($\mathbf{F}\mathbf{L}\mathbf{F}^* = \mathbf{L}$ since diagonal matrices are commutative) and as $\mathbf{U}_2^*\mathbf{B}\mathbf{U}_2 = \mathbf{T}_1\mathbf{L}\mathbf{T}_1^{-1}$ by (3) we have $\mathbf{U}_2^*\mathbf{B}\mathbf{U}_2 = \mathbf{U}_1\mathbf{A}\mathbf{U}_1^*$ or $\mathbf{B} = \mathbf{U}_3\mathbf{A}\mathbf{U}_3^*(\mathbf{U}_3 = \mathbf{U}_2\mathbf{U}_1)$ so that $\mathbf{B}$ and $\mathbf{A}$ are unitarily equivalent.

The diagonal elements of $\mathbf{X}^*\mathbf{X}$ are unity and since $\mathbf{X}^*\mathbf{X}$ is Hermitian it is determined by the $\frac{n(n-1)}{2}$ elements above the diagonal and as these elements are in general complex they involve $n(n-1)$ real constants. The element in the $p^{th}$ row and $q^{th}$ column of $\mathbf{Y}^*\mathbf{Y}$ is the product of the corresponding element in $\mathbf{X}^*\mathbf{X}$ by $\tau_q/\tau_p$ and on eliminating these $(n-1)$ ratios there remain $n(n-1) - (n-1)$ real unitary invariants. Included among these are the moduli of the elements of $\mathbf{X}^*\mathbf{X}$; denoting the element in the $p^{th}$ row and $q^{th}$ column of $\mathbf{X}^*\mathbf{X}$ by $(pq)$ the other unitary invariants are of the type $(pq)(qs)/(ps)$. Since the canonical forms of two unitarily equivalent matrices are unitarily equivalent and since the transforming matrix that sends one into another must be a diagonal phase matrix (since the common diagonal elements are assumed all different) the moduli of the various elements $c_q^p$ of the canonical form are unitary invariants as are also the various ratios $c_q^p c_s^q / c_s^p$.[3]

The discussion is more complicated when the matrix $\mathbf{A}$ is special. There are two distinctly different cases. One occurs when each multiple characteristic number has corresponding to it as many linearly distinct characteristic vectors as its degree of multiplicity (i.e., the elementary divisors of $\mathbf{A}$ are simple). To each multiple characteristic number $\lambda_p$ of multiplicity $q$ corresponds a linear space $S_q$ of characteristic vectors and we may choose as our characteristic vectors out of the subspace a set of $q$ mutually perpendicular unit vectors. The indeterminateness in the characteristic matrix $\mathbf{X}$ is now a phase matrix $\mathbf{F}$ which is not diagonal but consists of diagonal blocks, each of which is a unitary matrix. For example, if $n = 4$ and there are two simple characteristic numbers and one of multiplicity two $\mathbf{F}$ is of the type

$$\mathbf{F} = \begin{pmatrix} u_1^1 & 0 & 0 & 0 \\ 0 & u_2^2 & 0 & 0 \\ 0 & 0 & u_3^3 & u_4^3 \\ 0 & 0 & u_3^4 & u_4^4 \end{pmatrix}$$

(the letter $u$ being used to recall that $\mathbf{F}$ is unitary). The rest of the argument is as before, the only point of difference being the proof that $\mathbf{F}\mathbf{L}\mathbf{F}^* = \mathbf{L}$; this follows on account of the equalities between the characteristic numbers, i.e., the diagonal elements of $\mathbf{L}$. The number of unitary invariants is lessened; it will be sufficient to give the results for a single

case, say $n = 3$, where there is one characteristic number of multiplicity two. The norm $\mathbf{X}^*\mathbf{X}$ is of the form

$$\mathbf{X}^*\mathbf{X} = \begin{pmatrix} \overline{1} & (\overline{12}) & (\overline{13}) \\ (12) & 1 & 0 \\ (13) & 0 & 1 \end{pmatrix}$$

and $\mathbf{F} = \begin{pmatrix} \tau & 0 & 0 \\ 0 & u_2^2 & u_3^2 \\ 0 & u_2^3 & u_3^3 \end{pmatrix}$. A simple calculation shows that $(12)(\overline{12}) + (13)(\overline{13})$ is the only new unitary invariant.

The most complicated case of all occurs when the matrix $\mathbf{A}$ is degenerate, i.e., when the number of independent characteristic vectors is less than $n$. To outline the procedure let us consider a matrix $\mathbf{A}$ of order three with one simple characteristic number $\lambda_1$ and one double characteristic number $\lambda_2$ also of index one. The two characteristic vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ (of unit magnitude) are determinate save for scalar factors of modulus unity. A third vector $\mathbf{x}_3$ is determined by the equation $\mathbf{A}\mathbf{x}_3 = \mathbf{x}_2 + \lambda_2\mathbf{x}_3$ and by the fact of being perpendicular to $\mathbf{x}_2$. Denoting by $\mathbf{X}$ the matrix whose columns are the vectors $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ we have $\mathbf{AX} = \mathbf{X}\Lambda$, where $\Lambda$ is the matrix.

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

As before the factorization $\mathbf{X} = \mathbf{UT}$ gives Schur's reduction to the canonical triangular form. The vector $\mathbf{x}_3$ has now the same indeterminateness as $\mathbf{x}_2$ and so the phase matrix $\mathbf{F}$ is of the special type

$$\mathbf{F} = \begin{pmatrix} \tau_1 & 0 & 0 \\ 0 & \tau_2 & 0 \\ 0 & 0 & \tau_2 \end{pmatrix}$$

so that $\mathbf{F}\Lambda\mathbf{F}^* = \Lambda$. The norm $\mathbf{X}^*\mathbf{X}$ is again of the form

$$\begin{pmatrix} 1 & (\overline{12}) & (\overline{13}) \\ (12) & 1 & 0 \\ (13) & 0 & (33) \end{pmatrix}$$

and distinguishing the corresponding quantities referring to the second matrix by primes we have

$$(12)' = (12)\frac{\tau_1}{\tau_2}; \quad (13)' = (13)\frac{\tau_1}{\tau_2}; \quad (33)' = (33)$$

so that we have the three unitary invariants $(12)(\overline{12})$; $(12)/(13)$ and $(33)$.

[1] For a description of this process and an explicit expression of the elements of **T**, see A. Wintner, *Spektraltheorie der unendlichen matrizen*, 1929, pp. 32–61.

[2] These PROCEEDINGS, **17**, 676–678 (1931).

[3] This fact was pointed out to me by Dr. John Williamson of this department. This particular form of statement is convenient in a solution of the problem of determining the region of the complex plane covered by the values of the composite form $(\mathbf{Ax}|\mathbf{x})$ under the condition $(\mathbf{x}|\mathbf{x}) = 1$. This region, which is known to be convex, is bounded by an algebraic curve as will be shown in a forthcoming note.

---

# *CONFORMAL TENSORS*

## (*Second Note*)

BY TRACY YERKES THOMAS

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated December 18, 1931

1. In Note I we emphasized the fundamental idea of covariant differentiation as the construction of conditions of integrability in tensorial form of the equations of transformation of the components of a given tensor $T$. This suggests that we replace these latter equations, in case $K = 0$ or $L = 0$ so that the method in Note I does not apply, by an algebraically equivalent system of tensor character before performing the operation of covariant differentiation. We shall consider in this paragraph several types of equivalent systems in the above sense.

Denote by $T$ an arbitrarily given tensor of weight $W$ having the components $T^{p\ldots q}_{r\ldots s}$ and consider the conformal tensor with components

$$T^{p\ldots q}_{r\ldots s}\,T^{c\ldots d}_{e\ldots f} \equiv T^{p\ldots q\ c\ldots d}_{r\ldots s\ e\ldots f}, \tag{1.1}$$

this latter tensor is called the *square of the tensor* $T$ and will be denoted by the symbol $T^2$. In the equations of transformation of the components of $T^2$, namely,

$$\overline{T}^{i\ldots j\ g\ldots h}_{k\ldots l\ v\ldots w}\,u^p_i \ldots u^d_h = \left|u^a_b\right|^{2W} T^{p\ldots q\ c\ldots d}_{r\ldots s\ e\ldots f}\,u^r_k \ldots u^f_w, \tag{1.2}$$

let us take the indices $k\ldots l$ to be identical with the indices $v\ldots w$ and the indices $p\ldots q$ as identical with the indices $c\ldots d$. Then (1.2) reduces to

$$\overline{T}^{i\ldots j}_{k\ldots l}\,u^p_i \ldots u^q_j = \pm\left[\left|u^a_b\right|^W T^{p\ldots q}_{r\ldots s}\,u^r_k \ldots u^s_l\right],$$

and these equations are identical when the plus sign is selected in the right members, with the equations of transformation of the components

of the tensor $T$. Equations (1.2) therefore furnish an example of a system algebraically equivalent to the equations of transformation of the components of the tensor $T$.

If we define the tensor $T^m$ as the $m^{th}$ power of the tensor $T$ by equations analogous to (1.1) then it is evident that the equations of transformation of the components of $T^m$ are likewise algebraically equivalent to the equations of transformation of the components of the tensor $T$; in fact, for $m$ odd, even the above ambiguity of algebraic sign does not appear. It is evident therefore that the equations of transformation of $T^m$ can be used in place of the equations of transformation of the components of $T$ in the process of forming the conditions of integrability of these latter equations, i.e., the process of covariant differentiation.

Another type of algebraically equivalent system can be obtained by observing that the constants $V_i$ defined by

$$V_x \equiv 0, \quad V_\infty \equiv 1$$

constitute the components of a conformal vector of weight $-\frac{2}{n+2}$, i.e., these quantities undergo the transformation

$$\overline{V}_k = \left| u_b^a \right|^{-\frac{2}{n+2}} V_i u_k^i.$$

Now define

$$T_{r\ldots st}^{p\ldots q} \equiv V_i T_{r\ldots s}^{p\ldots q}, \tag{1.3}$$

so that

$$\overline{T}_{k\ldots lf}^{i\ldots j} u_i^p \ldots u_j^q = \left| u_b^a \right|^{W-\frac{2}{n+2}} T_{r\ldots st}^{p\ldots q} u_k^r \ldots u_l^s u_f^t. \tag{1.4}$$

If $f \neq \infty$ in (1.4) both members of these equations vanish identically, while if $f = \infty$ these equations become identical with the equations of transformation of the components of the tensor $T$. Hence (1.4) can be used in place of the equations of transformation of the components of the tensor $T$ in the process of constructing the conditions of integrability of these latter equations.

We will refer to the tensor with components given by (1.3) as the *augmented tensor* $T$. The vector with components $V_i$ will be called the *relative numerical vector* $V$. We can also define an *absolute numerical vector* $V$ by the equations

$$V^i = P^{ik} V_k, \tag{1.5}$$

where the quantities $P^{ik}$ have been previously defined. Then $V^\infty \equiv$

$V^i \equiv 0$ $(i = 1, \ldots, n)$, also $V^0 \equiv -1$, and these components satisfy the relations

$$V^i = \overline{V}^k u^i_k.$$

We see that these latter equations will hold for any value of the component $V^0$, but if this component is taken to be different from $-1$ the relations (1.5) will not be satisfied. It is obvious that a tensor might be defined, analogous to the above augmented tensor $T$, by using the absolute numerical vector $V$ in place of the relative numerical vector $V$.

2. *Case* $K = 0$.—The process defined in §2 of Note I for constructing the complete covariant derivative of the tensor $T$ with components $T^{p\ldots q}_{r\ldots s}$ is applicable provided that the number $K$ given by

$$K = 2M - 2N + 2(n+2)W + 2 - n$$

is different from zero; here $M$ denotes the number of covariant indices $r\ldots s$, $N$ the number of contravariant indices $p\ldots q$, $W$ the weight of the tensor $T$ and $n$ the dimensionality number of the space. If, now, $K = 0$ for the above tensor $T$, then $K = -2$ for the augmentated tensor $T$. Hence the complete covariant derivative of the augmentated tensor $T$ can be formed.

Using the formulae of Note I it follows readily that

$$T^{p\ldots q}_{r\ldots s\infty,\mu} \equiv T^{p\ldots q}_{r\ldots s,\mu}, \qquad (\mu = 1, \ldots, n) \tag{2.1}$$

$$T^{p\ldots q}_{r\ldots s\infty,0} \equiv \left[\frac{M - N + (n+2)W - 1}{n}\right] T^{p\ldots q}_{r\ldots s}, \tag{2.2}$$

$$T^{p\ldots q}_{r\ldots s\mu,\nu} \equiv \frac{1}{n}\, T^{p\ldots q}_{r\ldots s}\, G_{\mu\nu}, \qquad (\mu,\nu = 1, \ldots, n) \tag{2.3}$$

$$T^{p\ldots q}_{r\ldots s0,\nu} \equiv T^{p\ldots q}_{r\ldots s\mu,0} \equiv 0, \qquad (\mu, \nu = 1, \ldots, n) \tag{2.4}$$

where the left members of these equations represent components of the covariant derivative of the augmented tensor $T$. Now consider certain of the equations of transformation of the components of this latter tensor, namely, the following

$$\overline{T}^{i\ldots j}_{k\ldots l\infty,\,\alpha}\, u^p_i \ldots u^q_j = \left|u^a_b\right|^{W-\frac{2}{n+2}} T^{p\ldots q}_{r\ldots sl,\beta}\, u^r_k \ldots u^s_l u^t_\infty u^\beta_\alpha. \tag{2.5}$$

Except for the inconsequential fact that both members of (2.5) vanish identically for $\alpha = 0$, $n = 4$, these equations are exactly equivalent to the equations of transformation of the components of the incomplete covariant derivative of the tensor $T$. On this account, as well as on account of the equations (2.1) and (2.2), the process of forming the co-

variant derivative of the augmented tensor $T$ is closely analogous to the process of completing the incomplete covariant derivative of the tensor $T$ as defined in Note I, and for that reason can be looked upon as a new method of completing this latter tensor. We will therefore refer to the above covariant derivative of the augmented tensor $T$ as the *complete covariant derivative, $K = 0$, or simply as the covariant derivative, $K = 0$, of the tensor $T$.* Since $K = 0$ likewise for this latter tensor the sequence of the components of the successive covariant derivatives of the tensor $T$, has the form

$$T^{p\ldots q}_{r\ldots s,w_1w_2};\quad T^{p\ldots q}_{r\ldots s,w_1\ldots w_4};\quad T^{p\ldots q}_{r\ldots s,w_1\ldots w_6};\quad \ldots.$$

The number of covariant indices in the symbols of these components increases by jumps of 2 and the weights of the corresponding tensors by amounts of $-\frac{2}{n+2}$ as we proceed from left to right along the sequence. It follows from what was said in § 1 as well as from what was just said regarding equations (2.5) that the equations of transformation of the components in the above sequence express the conditions of integrability obtained by successive differentiation of the equations of transformation of the components of the original tensor $T$.

If $K = 0$ for the tensor $T$, then the integer $K$ is equal to $(n-2)(m-1)$ for the tensor $T^m$ defined in the preceding paragraph. Hence, for $m \gtreqless 2$, the number $K > 0$ for the tensor $T^m$ and we can construct the complete covariant derivative of this tensor; since, moreover, the effect of the covariant differentiation is to increase the number $K$ by 2, it is seen that $K$ will never equal zero in the process of forming the successive covariant derivatives of $T^m$.

3. *Case $L = 0$.*—If $L = 0$ for the incomplete conformal tensor $D$ defined in § 3 of Note I, then $L = -1$ for the *incomplete augumented tensor $D$* defined by the components

$$D^{p\ldots q}_{r\ldots s1\mu\nu} \equiv V_t\, D^{p\ldots q}_{r\ldots s\mu\nu}.$$

This latter tensor can therefore be completed by the method in § 3 of Note I and the value of the integer $K$ is equal to $n - 4$ for the *complete augumented tensor $D$* so obtained. Hence if $n = 4$ we must construct covariant derivatives, $K = 0$, of the augmented tensor $D$; otherwise the process of covariant differentiation of Note I applies.

4. *Application to the Conformal Curvature Tensor (n Even).*—For even values of the dimensionality number $n$ the method of § 2 and § 3 leads immediately to the following sequences for the complete conformal curvature tensor and its successive covariant derivatives.

$$
\begin{array}{llll}
(n = 4) & B^i_{jklm}; & B^i_{jklmw_1w_2}; & B^i_{jklmw_1 \ldots w_4}; \quad \ldots \\
(n = 6) & B^i_{jkl}; & B^i_{jklw_1w_2}; & B^i_{jklw_1 \ldots w_4}; \quad \ldots \\
(n = 8) & B^i_{jkl}; & B^i_{jklm}; & B^i_{jklmw_1w_2}; \quad \ldots \\
 & \cdot & \cdot & \cdot \\
 & \cdot & \cdot & \cdot \\
 & \cdot & \cdot & \cdot
\end{array}
$$

Each of these sequences begins with the complete conformal curvature tensor; and those tensors whose components appear along the main diagonal of the above infinite matrix of components, are each of weight $-\frac{2}{n+2}$. It is evident that the tensors whose components appear in the above sequences, when combined with the fundamental conformal tensor $G$, constitute a complete set of invariants of the conformal Riemann space.

---

## *INTRANSITIVE GROUPS OF MOTIONS*

By Luther Pfahler Eisenhart

Department of Mathematics, Princeton University

Communicated December 28, 1931

The continuous groups of motions of a Riemannian manifold have been the subject of study of geometers for many years, and more recently their rôle in theoretical physics has been revealed. As a basis for the classification of such groups Fubini[1] announced the theorem:

*If a $G_r$ is an intransitive group of motions of a $V_n$ and the minimum invariant varieties are of dimensionality $q$, a coördinate system can be found in which $G_r$ is a transitive group on $q$ variables.* The proof of this theorem given by Fubini is satisfactory for the case when $q = n - 1$, but it is not convincing for $q < n - 1$. It is the purpose of this note to give a proof of this theorem for spaces with a definite quadratic form (the case considered by Fubini) and to give sufficient conditions for it in the case when the quadratic form is indefinite. Throughout the paper the variables are understood to be real.

1. *Equations of Continuous Groups.*—We consider a continuous group of order $r$, say $G_r$, expressed in terms of $n$ coördinates $x^1, \ldots, x^n$, and denote by

$$X_a f = \xi^\alpha_a \frac{\partial f}{\partial x^\alpha} \qquad \begin{pmatrix} a = 1, \ldots, r; \\ \alpha = 1, \ldots, n \end{pmatrix} \qquad (1.1)$$

the symbols of the group. Here, and in what follows, a repeated index in a term (in this case $\alpha$) indicates that the term is the sum of terms as the index takes on all its values. The fundamental theorem of continuous groups is that a necessary and sufficient condition that $X_a f$ are the generators of a group is that

$$(X_a, X_b)f = c_{ab}{}^e X_e f \qquad (a, b, e = 1, \ldots, r) \tag{1.2}$$

where the $c$'s are constants satisfying the conditions

$$c_{ab}{}^e = -c_{ba}{}^e, \tag{1.3}$$

$$c_{ab}{}^e c_{ed}{}^f + c_{bd}{}^e c_{ea}{}^f + c_{da}{}^e c_{eb}{}^f = 0. \tag{1.4}$$

Substituting from (1.1) in (1.2), we obtain

$$\xi_a^\alpha \frac{\partial \xi_b^\beta}{\partial x^\alpha} - \xi_b^\alpha \frac{\partial \xi_a^\beta}{\partial x^\alpha} = c_{ab}{}^e \xi_e^\beta. \tag{1.5}$$

We denote by $q$ the generic rank of the matrix

$$M = \| \xi_a^\alpha \|, \tag{1.6}$$

that is, the rank for general values of the $x$'s. Evidently $q \geq r$. If $q = n$, the group is transitive; if $q < n$, intransitive. If $q < r$, we may without loss of generality assume that the matrix $\| \xi_h^\alpha \|$, for $h = 1, \ldots, q$, is of rank $q$, and put

$$\xi_p^\alpha = \varphi_p^h \xi_h^\alpha \qquad \begin{pmatrix} h = 1, \ldots, q; \\ p = q + 1, \ldots, r \end{pmatrix} \tag{1.7}$$

where the quantities $\varphi_p^h$ are functions of the $x$'s. Then (1.5) may be written

$$\xi_a^\alpha \frac{\partial \xi_b^\beta}{\partial x^\alpha} - \xi_b^\alpha \frac{\partial \xi_a^\beta}{\partial x^\alpha} = (c_{ab}{}^h + c_{ab}{}^p \varphi_p^h) \xi_h^\beta. \tag{1.8}$$

If we take $b = p$ in (1.8) and replace $\xi_p^\beta$ by its expression (1.7), the result is equivalent, in consequence of equations of the form (1.8) in $a$ and $h$, to

$$X_a \varphi_p^h = \Phi_{ap}^h, \tag{1.9}$$

where

$$\Phi_{ap}^h = c_{ap}{}^h + c_{ap}{}^s \varphi_s^h - \varphi_p^i (c_{ai}{}^h + c_{ai}{}^s \varphi_s^h) \qquad \begin{pmatrix} a = 1, \ldots, r; \\ h, i = 1, \ldots, q; \\ p, s = q + 1, \ldots, r \end{pmatrix} \tag{1.10}$$

When $q < n$, that is when $G_r$ is intransitive, the equations $X_a f = 0$ form a completely integrable system, and because of (1.7) there are $q$

independent equations in the set, and consequently there are $n - q$ independent solutions $\varphi^\sigma(x)$. Without loss of generality, we may assume that the determinant $\left|\frac{\partial\varphi^\sigma}{\partial x^\tau}\right|$ for $\tau = q+1, \ldots, n$ is not zero. If we effect the transformation

$$x'^\lambda = \chi^\lambda,\ x^\sigma = \varphi^\sigma(x) \begin{pmatrix} \lambda = 1, \ldots, q; \\ \sigma = q+1, \ldots, n \end{pmatrix}, \tag{1.11}$$

and note that the functions $\xi_a^\alpha$ transform as the components of a contravariant vector, that is,

$$\xi'^\alpha_a = \xi^\beta_a \frac{\partial x'^\alpha}{\partial x^\beta}, \tag{1.12}$$

then in the new coördinate system, which we now call $x^\alpha$ (dropping primes), we have

$$\xi^\sigma_a = 0 \begin{pmatrix} a = 1, \ldots, r; \\ \sigma = q+1, \ldots, n \end{pmatrix}. \tag{1.13}$$

Since in the new coördinate system the matrix $\| \xi^\lambda_h \|$ for $h = 1, \ldots, q$ is of rank $q$, a set of functions $\xi_{h\lambda}$ are uniquely defined by

$$\xi_{h\lambda}\xi^\mu_h = \delta^\mu_\lambda,\ \xi_{h\lambda}\xi^\lambda_i = \delta_{ih} \qquad (\lambda, \mu, i, h = 1, \ldots, q), \tag{1.14}$$

where

$$\delta^\mu_\lambda = 1 \text{ or } 0, \text{ as } \mu = \lambda \text{ or } \mu \neq \lambda$$
$$\delta_{ih} = 1 \text{ or } 0, \text{ as } i = h \text{ or } i \neq h.$$

In this coördinate system equations (1.8), when $a$ and $b$ take values 1 to $q$ reduce to

$$\xi^\lambda_h \frac{\partial \xi^\mu_i}{\partial x^\lambda} - \xi^\lambda_i \frac{\partial \xi^\mu_h}{\partial x^\lambda} = (c^j_{hi} + c^p_{hi}\,\varphi^j_p)\xi^\mu_j \qquad \begin{pmatrix} (\lambda, \mu, h, i, j = 1, \ldots, q; \\ p = q+1, \ldots, r \end{pmatrix}, \tag{1.15}$$

and equations (1.9) for $a = 1, \ldots, q$ may be written in the form

$$\frac{\partial \varphi^h_p}{\partial x^\lambda} = \Phi^h_{ip}\,\xi_{i\lambda}, \tag{1.16}$$

in consequence of (1.14).

We define functions $L^\alpha_{\beta\mu}$ by

$$L^\lambda_{\beta\mu} = \xi^\lambda_h \frac{\partial \xi_{h\mu}}{\partial x^\beta} = -\xi_{h\mu} \frac{\partial \xi^\lambda_h}{\partial x^\beta},\ L^\sigma_{\beta\mu} = 0 \begin{pmatrix} \lambda, \mu = 1, \ldots, q; \\ \beta = 1, \ldots, n; \\ \sigma = q+1, \ldots, n \end{pmatrix}, \tag{1.17}$$

(The reason for this definition will appear in §2.) From (1.17) we have, in consequence of (1.14)

$$\frac{\partial \xi_h^\lambda}{\partial x^\beta} + \xi_h^\mu L_{\beta\mu}^\lambda = 0 \tag{1.18}$$

and

$$\frac{\partial \xi_{h\lambda}}{\partial x^\beta} - \xi_{h\mu} L_{\beta\lambda}^\mu = 0. \tag{1.19}$$

From (1.17) and (1.15) we have

$$L_{\mu\nu}^\lambda - L_{\nu\mu}^\lambda = (c_{hi}^j + c_{hi}^p \varphi_p^j)\xi_{h\nu}\xi_{i\mu}\xi_j^\lambda. \tag{1.20}$$

Since the determinant $|\xi_h^\lambda|$ is different from zero, the conditions of integrability of (1.18) are

$$\frac{\partial L_{\alpha\mu}^\lambda}{\partial x_\beta} - \frac{\delta L_{\beta\mu}^\lambda}{\partial x^\alpha} + L_{\alpha\mu}^\pi L_{\beta\pi}^\lambda - L_{\beta\mu}^\pi L_{\alpha\pi}^\lambda = 0 \quad \begin{pmatrix} \alpha, \beta = 1, \ldots, n; \\ \lambda, \mu, \pi = 1, \ldots, q \end{pmatrix}. \tag{1.21}$$

These are necessarily identities, as may be verified by substitution from (1.17).

If we replace $\beta$ in (1.21) by $\nu$ for $\nu = 1, \ldots, q$, and subtract from these equations the corresponding ones obtained by interchanging $\mu$ and $\nu$, we obtain

$$\frac{\partial L_{\alpha\mu}^\lambda}{\partial x^\nu} - \frac{\partial L_{\alpha\nu}^\lambda}{\partial x^\mu} = \frac{\partial}{\partial x^\alpha}(L_{\nu\mu}^\lambda - L_{\mu\nu}^\lambda) + L_{\alpha\pi}^\lambda(L_{\nu\mu}^\pi - L_{\mu\nu}^\pi) - L_{\alpha\mu}^\pi L_{\nu\pi}^\lambda + L_{\alpha\nu}^\pi L_{\mu\pi}^\lambda.$$

If we put

$$L_{\alpha\mu\nu}^\lambda = \frac{\partial L_{\alpha\nu}^\lambda}{\partial x^\mu} - \frac{\partial L_{\alpha\mu}^\lambda}{\partial x^\nu} + L_{\alpha\nu}^\pi L_{\pi\mu}^\lambda - L_{\alpha\mu}^\pi L_{\pi\nu}^\lambda, \tag{1.22}$$

the preceding equations may be written

$$L_{\alpha\mu\nu}^\lambda = \frac{\partial}{\partial x^\alpha}(L_{\mu\nu}^\lambda - L_{\nu\mu}^\lambda) + L_{\alpha\pi}^\lambda(L_{\mu\nu}^\pi - L_{\nu\mu}^\pi) + L_{\alpha\nu}^\pi(L_{\pi\mu}^\lambda - L_{\mu\pi}^\lambda) - L_{\alpha\mu}^\pi(L_{\pi\nu}^\lambda - L_{\nu\pi}^\lambda).$$

In consequence of (1.20), (1.18) and (1.19) these equations are reducible to

$$L_{\alpha\mu\nu}^\lambda = c_{hi}^p \frac{\partial \varphi_p^j}{\partial x^\alpha} \xi_{h\nu}\xi_{i\mu}\xi_j^\lambda. \tag{1.23}$$

We inquire whether there exists a transformation of coördinates of the form

$$x^\lambda = \varphi^\lambda(x'^1, \ldots, x'^n), x^\sigma = x'^\sigma \quad \begin{pmatrix} \lambda = 1, \ldots, q; \\ \sigma = q+1, \ldots, n \end{pmatrix} \tag{1.24}$$

such that in the new coördinate system $\xi'^{\lambda}_{h}$ are independent of $x'^{\sigma}$ for $\sigma = q + 1, \ldots, n$. The inverse of (1.24) is of the form

$$x'^{\lambda} = \bar{\varphi}^{\lambda}(x^1, \ldots, x^n), \quad x'^{\sigma} = x^{\sigma}. \tag{1.25}$$

Then because of (1.13) we have $\xi'^{\sigma}_{h} = 0$. If the desired coördinate system exists, on differentiating with respect to $x'^{\sigma}$ the first set of the equations

$$\xi^{\mu}_{h} = \xi'^{\lambda}_{h} \frac{\partial x^{\mu}}{\partial x'^{\lambda}}, \quad \xi'^{\lambda}_{h} = \xi^{\mu}_{h} \frac{\partial x'^{\lambda}}{\partial x^{\mu}}, \tag{1.26}$$

and making use of (1.18), the second set of (1.26) and the fact that the determinant $|\xi^{\lambda}_{h}|$ is not zero, we obtain the following equations of condition:

$$\frac{\partial x'^{\nu}}{\partial x^{\lambda}} \frac{\partial^2 x^{\mu}}{\partial x'^{\nu} \partial x'^{\sigma}} + L^{\mu}_{\alpha\lambda} \frac{\partial x^{\alpha}}{\partial x'^{\sigma}} = 0. \tag{1.27}$$

For the transformation (1.24) we have

$$\frac{\partial x'^{\nu}}{\partial x^{\lambda}} \frac{\partial x^{\lambda}}{\partial x'^{\mu}} = \delta^{\nu}_{\mu}.$$

Hence if we multiply (1.27) by $\dfrac{\partial x^{\lambda}}{\partial x'^{\pi}}$ and sum for $\lambda$, we obtain, with change of indices,

$$\frac{\partial^2 x^{\mu}}{\partial x'^{\nu} \partial x'^{\sigma}} + L^{\mu}_{\alpha\lambda} \frac{\partial x^{\alpha}}{\partial x'^{\sigma}} \frac{\partial x^{\lambda}}{\partial x'^{\nu}} = 0. \tag{1.28}$$

When we differentiate this equation with respect to $x'^{\pi}$ for $\pi = 1, \ldots, q$, reduce the resulting equation by means of (1.28) and subtract from the result the equation obtained by interchanging $\nu$ and $\pi$, we obtain

$$\frac{\partial x^{\alpha}}{\partial x'^{\sigma}} \frac{\partial x^{\lambda}}{\partial x'^{\nu}} \frac{\partial x^{\theta}}{\partial x'^{\pi}} L^{\mu}_{\alpha\lambda\theta} = 0. \tag{1.29}$$

When we differentiate equation (1.28) with respect to $x'^{\tau}$ for $\tau = q + 1, \ldots, n$, reduce the resulting equation by means of (1.28) and subtract from the result the equation obtained by interchanging $\sigma$ and $\tau$, we obtain an equation which vanishes identically because of (1.21).

2. *Groups of Motions.*—If the variables $x^{\alpha}$ are interpreted as the coördinates of points of an $n$ dimensional space, $V_n$, the space is Riemannian when a metric is assigned by a quadratic differential form $g_{\alpha\beta}dx^{\alpha}dx^{\beta}$ where without loss of generality it may be assumed that $g_{\alpha\beta} = g_{\beta\alpha}$, and where the determinant $|g_{\alpha\beta}|$ is not zero. In order that the space admit a given $G_r$ as a group of motions, that is for each infinitesimal transformation of the group $\delta(g_{\alpha\beta}dx^{\alpha}dx^{\beta}) = 0$, it is necessary and sufficient that the following equations of Killing hold in each coördinate system:[2]

$$\xi_a^\alpha \frac{\partial g_{\beta\gamma}}{\partial x^\alpha} + g_{\beta\alpha} \frac{\partial \xi_a^\alpha}{\partial x^\gamma} + g_{\gamma\alpha} \frac{\partial \xi_a^\alpha}{\partial x^\beta} = 0 \begin{pmatrix} \alpha, \beta, \gamma = 1, \ldots, n; \\ a = 1, \ldots, r \end{pmatrix}. \quad (2.1)$$

If $q < n$, then in the special coördinate system for which (1.13) hold, these equations become for $a = 1, \ldots, q$

$$\xi_h^\lambda \frac{\partial g_{\beta\gamma}}{\partial x^\lambda} + g_{\beta\lambda} \frac{\partial \xi_h^\lambda}{\partial x_\gamma} + g_{\gamma\lambda} \frac{\partial \xi_h^\lambda}{\partial x^\beta} = 0. \quad \begin{pmatrix} \lambda, h = 1, \ldots, q; \\ \beta, \gamma = 1, \ldots, n \end{pmatrix} \quad (2.2)$$

If these equations are multiplied by $\xi_{h\mu}$ and $h$ is summed, we have, in consequence of (1.14) and (1.18), the equivalent set of equations

$$\frac{\partial g_{\beta\gamma}}{\partial x^\mu} = g_{\beta\lambda} L^\lambda_{\gamma\mu} + g_{\gamma\lambda} L^\lambda_{\beta\mu}. \quad (2.3)$$

When $q = r$, equations (2.2) are the only conditions. However, when $q < r$, we have in addition (2.1) when $a$ takes the values $q + 1$ to $r$. When in these equations we substitute from (1.7) and (1.13), we obtain with the aid of (2.3) the following equivalent system:

$$\xi_h^\lambda \left( g_{\beta\lambda} \frac{\partial \varphi_p^h}{\partial x_\gamma} + g_{\lambda\gamma} \frac{\partial \varphi_p^h}{\partial x^\beta} \right) = 0. \quad \begin{pmatrix} \lambda, h = 1, \ldots, q; \\ p = q + 1, \ldots, r; \\ \beta, \gamma = 1, \ldots, n \end{pmatrix} \quad (2.4)$$

Thus when $q < r$, the systems (2.2) and (2.4) are the conditions of the problem.

Expressing the condition of integrability of equations (2.3), we obtain

$$g_{\beta\lambda} L^\lambda_{\gamma\mu\nu} + g_{\gamma\lambda} L^\lambda_{\beta\mu\nu} = 0,$$

where $L^\lambda_{\gamma\mu\nu}$ is defined by (1.22). When $q = r$ and $r \gtrless n$, it follows from (1.23) that $L^\lambda_{\beta\mu\nu} = 0$. Hence, equations (2.3), which are the only conditions to be satisfied in this case, form a completely integrable system. Hence a solution is determined by arbitrary initial values of the $g$'s, which when $n > r$ may be arbitrary functions of the variables $x^{r+1}, \ldots, x^n$. Hence we have the theorem:[8]

*Any r-parameter continuous group on n variables such that the generic rank of the matrix* $\| \xi_a^\alpha \|$ *is* $r(\gtrless n)$ *is the group of motions of a Riemannian manifold whose fundamental tensor* $g_{\alpha\beta}$ *involves* $n(n+1)/2$ *arbitrary functions of* $n - r$ *variables.*

When $n = r$, that is when $G_r$ is simply transitive, there are $n(n+1)/2$ arbitrary constants; the theorem for this particular case was established by Bianchi.

It is known that when a group is Abelian, that is, when all of the constants of composition are zero, the coördinates can be chosen so that

$$\xi_h^\lambda = \delta_h^\lambda, \ \xi_h^\sigma = 0 \begin{pmatrix} \lambda, h = 1, \ldots, q \\ \sigma = q+1, \ldots, n \end{pmatrix}.$$

In this case the functions $\Phi_{\alpha\beta}^h$ (1.10) are zero and from (1.16) it follows that the quantities $\varphi_p^h$ are independent of $x^1, \ldots, x^q$. Then from (2.4) we have $g_{\lambda\mu} \dfrac{\partial \varphi_p^\mu}{\partial x^\sigma} = 0$, for $\sigma = q+1, \ldots, n$. Hence the $\varphi$'s are constants unless the determinant is zero, which is impossible if the fundamental quadratic form is definite.[4] Consequently only when $q = r$ is an intransitive Abelian group a group of motions. Moreover, every such Abelian group is a group of motions, and the quantities $g_{\alpha\beta}$ are arbitrary functions of $x^{q+1}, \ldots, x^n$, as follows from (2.3), since the quantities $L_{\alpha\mu}^\lambda$ are zero in this case. When the group is Abelian and simply transitive, the $g_{\alpha\beta}$ are constants.[5]

Returning to the case of non-Abelian groups for which $q = r$, we observe from (1.23) that the quantities $L_{\beta\mu\nu}^\lambda$ are zero and from (1.29) that the system of equations (1.28) are completely integrable. Consequently it is possible to express the functions $\varphi^\lambda$ in (1.24) as a power series by means of (1.28). In order to insure that the transformation be non-singular, the initial values of $\dfrac{\partial x^\lambda}{\partial x'^\nu}$ must be chosen so that their determinant is not equal to zero. The initial values of $\dfrac{\partial x^\lambda}{\partial x'^\sigma}$ may be taken arbitrarily, as also the derivatives of $x^\mu$ of any order with respect to $x^1, \ldots, x^q$, or with respect to $x^{q+1}, \ldots, x^n$. Hence we have the theorem:

*When the rank of the matrix* $\| \xi_\alpha^\alpha \|$ *of an intransitive group* $G_r$ *is* $r$, *there exists a coördinate system for the* $V_n$ *in which the* $\xi$'*s are functions of* $x^1, \ldots, x^r$ *at most, the minimum invariant varieties being defined by* $x^{r+1} = const., \ldots, x^n = const.$

3. *When* $q < r$.—Since by hypothesis the determinant $|g_{\alpha\beta}|$ is not zero, a set of functions $g^{\alpha\beta}$ is defined by

$$g^{\alpha\beta} g_{\beta\gamma} = \delta_\gamma^\alpha.$$

If we multiply (2.4) by $g^{\beta\sigma}$ for $\sigma = q+1, \ldots, n$ and sum for $\beta$, we have, because of (1.13),

$$\xi_h^\lambda g_{\gamma\lambda} g^{\beta\sigma} \frac{\partial \varphi_p^h}{\partial x^\beta} = 0.$$

If we multiply this equation by $g^{\gamma\mu}$ and sum for $\gamma$, and note that the determinant $|\xi_h^\lambda|$ is not zero, we have that the $q(n-q)$ functions $\varphi_p^h$ are solutions of the $n - q$ equations

$$g^{\beta\sigma}\frac{\partial\varphi}{\partial x^{\beta}} = 0 \quad \begin{pmatrix} \beta = 1, \ldots, n \\ \sigma = q+1, \ldots, n \end{pmatrix}. \tag{3.1}$$

Hence at most $q$ of the $\varphi$'s are independent.

Suppose that $q$ of them are independent, and denote them by $\varphi^1, \ldots, \varphi^q$; then the matrix $\left\| \frac{\partial\varphi^h}{\partial x^{\alpha}} \right\|$, for $h = 1, \ldots, q$; $\alpha = 1, \ldots, n$, is of rank $q$. If the determinant $\left| \frac{\partial\varphi^h}{\partial x^{\lambda}} \right|$ for $\lambda = 1, \ldots, q$ is zero, there exist functions $A_h$ such that

$$A_h \frac{\partial\varphi^h}{\partial x^{\lambda}} = 0. \tag{3.2}$$

Then from (3.1) we have

$$g^{\tau\sigma}A_h \frac{\partial\varphi^h}{\partial x^{\tau}} = 0 \qquad (\tau = q+1, \ldots, n) \tag{3.3}$$

Hence if $|g^{\tau\sigma}| \neq 0$, we have $A_h \frac{\partial\varphi^h}{\partial x^{\tau}} = 0$, which together with (3.2) imply that the $\varphi$'s are not independent. But $|g^{\sigma\tau}| \neq 0$, if the fundamental quadratic form is definite.[4] Consequently the determinant $\left| \frac{\partial\varphi^h}{\partial x^{\lambda}} \right|$ is not zero in this case, and accordingly a non-singular transformation of coördinates is defined by

$$x'^{\lambda} = \varphi^{\lambda}, \; x'^{\sigma} = x^{\sigma}. \tag{3.4}$$

In this new coördinate system (dropping primes) we have from the equations (1.16) for the functions $\varphi^{\lambda}$ equations of the form

$$\Phi_i^{\lambda}\xi_{i\mu} = \delta_{\mu}^{\lambda}, \tag{3.5}$$

where the $\Phi_i^{\lambda}$ are certain of the functions $\Phi_{ip}^{h}$ and being functions of $\varphi^{\lambda}$ are independent of $x^{q+1}, \ldots, x^n$. From these equations and (1.14) it follows that

$$\xi_h^{\lambda} = \Phi_h^{\lambda}. \tag{3.6}$$

From these equations and (1.7) it follows that $\xi_{\alpha}^{\lambda}$ are functions of $\varphi^{\lambda}$ and consequently of $x^1, \ldots, x^q$ at most. Hence we have:

*When $q$ of the functions $\varphi_p^h$ are independent and the fundamental quadratic form is definite, there exists a coördinate system for which (1.13) hold, and the components $\xi_a^{\lambda}$ are functions of $x^1, \ldots, x^q$ at most.*

From the preceding argument it follows also that:

*When the fundamental form is indefinite and the coördinate system is such that (1.13) hold and there are $q$ of the functions $\varphi_p^h$ of which the jacobian with*

*respect to* $x^1, \ldots, x^q$ *is not zero, then in the coördinate system* (3.4) *the components* $\xi'^{\lambda}_{a}$ *are functions of* $x^1, \ldots, x'^q$ *at most.*

We consider next the case when there are $s(< q)$ independent functions $\varphi^h_p$, denoted by $\varphi^1, \ldots, \varphi^s$. Suppose that the jacobian matrix of these $\varphi$'s with respect to $x^1, \ldots, x^q$ is less than $s$. Then we have equations of the form (3.2) and (3.3) in which $h$ takes the values 1 to $s$. Hence, unless the determinant $|g^{\sigma\tau}|$ is zero, we have from (3.3) that the rank of $\left\| \frac{\partial \varphi^h}{\partial x^\alpha} \right\|$ is less than $s$, contrary to hypothesis. If then the fundamental form is definite, there is no loss in generality in assuming that the determinant $\left| \frac{\partial \varphi^h}{\partial x^{\lambda_l}} \right|$ for $\lambda_1, h = 1, \ldots, s$ is not zero. When we effect the transformation of coördinates

$$x'^{\lambda_1} = \varphi^{\lambda_1}, \quad x'^{\lambda_2} = x^{\lambda_2}, \quad x'^{\sigma} = x^{\sigma} \begin{pmatrix} \lambda_1 = 1, \ldots, s; \\ \lambda_2 = s+1, \ldots, q; \\ \sigma = q+1, \ldots, n \end{pmatrix}$$

in the new coördinate system all the functions $\varphi^h_p$ are functions of $x^1, \ldots, x^s$.

Proceeding as above, we have in place of (3.6) $\xi^{\lambda_1}_h = \Phi^{\lambda_1}_h$, where the $\Phi$'s are functions of $x^1, \ldots, x^s$. We apply now a transformation of coordinates

$$x'^{\lambda_1} = x^{\lambda_1}, \quad x'^{\lambda_2} = \varphi^{\lambda_2}(x'^1, \ldots, x'^n), \quad x'^{\sigma} = x^{\sigma},$$

after the manner of (1.24). In place of (1.28) we have

$$\frac{\partial^2 x^{\lambda_2}}{\partial x'^{\tau} \partial x'^{\sigma}} + L^{\lambda_2}_{\alpha\mu} \frac{\partial x^\alpha}{\partial x'^{\sigma}} \frac{\partial x^\mu}{\partial x'^{\tau}} = 0. \tag{3.7}$$

From (1.23) it follows that $L^{\lambda}_{\alpha\mu\nu} = 0$ for $\alpha \gtrless = s+1, \ldots, n$. Consequently equations (3.7) are completely integrable as follows from (1.29) for the system (3.7). Hence the first of the above theorems holds when the number of independent functions $\varphi^h_p$ is less than $q$, and we have the theorem:

*When the fundamental quadratic form is definite there exists a coördinate system for which equations* (1.13) *hold and the components* $\xi^{\lambda}_{a}$, *for* $\lambda = 1, \ldots, q$ *are functions of* $x^1, \ldots, x^q$ *at most.*

Also we have:

*When the fundamental quadratic form is indefinite and there are* $s(\gtrless q)$ *independent functions* $\varphi^h_p$, *if for a coördinate system in which* (1.13) *hold, the jacobian matrix of the s independent functions with respect to* $x^\tau, \ldots, x_q$ *if of rank s, there exists a coördinate system in which the components* $\xi^{\lambda}_{\alpha}$ *are functions of* $x^1, \ldots, x^q$ *at most.*

[1] "Sugli spazii che ammettono un gruppo continuo di movimenti," *Annali di Matematica*, ser. 3, **8**, p. 40 (1903).

[2] Cf. Eisenhart, *Riemannian Geometry*, p. 234, 1926.

[3] Cf. Fubini, loc. cit., p. 54.

[4] If the determinant $|\ g_{\lambda\mu}\ |$ for $\lambda, \mu = 1, \ldots, q, (q < n)$ is zero, real quantities $\xi^\lambda$ are defined by the equations $g_{\lambda\mu}\xi^\lambda = 0$. The vector with the components $\xi^\lambda$, $\xi^\sigma = 0$ $(\sigma = q + 1, \ldots, n)$ is a real null vector, which is impossible, if the fundamental form is definite. In a similar manner it can be shown that the determinant $|\ g^{\sigma\tau}\ |$ for $\sigma,\tau = 1, \ldots, p$ cannot be zero for a definite form.

[5] Cf. Bianchi, *Lezioni sulla teoria dei gruppi continui finiti di trasformazioni*, p. 521, 1918.

---

## *ON CERTAIN PROPERTIES OF SEPARABLE SPACES*

BY SOLOMON LEFSCHETZ

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated January 14, 1932

1. Let $R$ be a separable metric space, $\Delta$ a *normal development* of $R$, that is a sequence of finite coverings by open sets $\{\Sigma^i\} = \{\{U^{i\alpha}\}\}$, with the following properties: $(a)$ the $U$'s constitute a fundamental set of neighborhoods: $(b)$ every $\overline{U}^{i+1}$ is on some $U^i$; $(c)$ if $\overline{U}^{i\alpha} \subset U^{j\beta}$, for $h$ high enough every $\overline{U}^{h\gamma}$ meeting $\overline{U}^{i\alpha}$ is on $U^{j\beta}$; $(d)$ every decreasing sequence $\{U^{N_i\alpha_i}\}, (\overline{U}^{Ni+1\alpha_i, \alpha_{i+1}} \subset U^{N_i\alpha_i})$ converges to a point or to zero. Developments of a very general type have been introduced by E. H. Moore, and many have been considered on repeated occasions in the theory of abstract spaces. In particular, spaces possessing a certain type called *regular* have been investigated with considerable success by Chittenden and Pitcher, notably in connection with the problem of metrization. Normal developments which are considered here for the first time appear to possess just the right degree of generality for the comfortable treatment of separable metric spaces. By means of them a number of the important theorems in dimension theory can be derived with little more trouble than in the compact case. Above all they have enabled us to obtain the proof as yet unavailable of a basic imbedding theorem on separable spaces.

2. The fundamental proposition is the

*Existence theorem. R always possesses a normal development.*

We shall only outline the proof here, reserving the details for publication elsewhere. Let us identify $R$ with its topological image on the Hilbert parallelotope $H$ image whose existence has been proved by Urysohn. Since $H$ is compact it possesses a finite covering by open sets $\sigma$, and the intersections of its sets with $R$ determine $\Sigma^1 = \{U^{1\alpha}\}$. We choose a metric such that there exists an $\eta > 0$ such that any subset of $R$ whose diameter $< \eta$ is on a $U^{1\alpha}$. This is done by modifying the mapping on $H$, the new distance function always yielding distances greater than the old. In

the same manner as above we select for $\Sigma^2$ an $\epsilon_2$-covering $\{U^{2\alpha}\}$, $\epsilon_2 < \eta$, etc. However, when we pass from $\Sigma^i$ to $\Sigma^{i+1}$ we have to choose a metric such that whenever pairs of closed sets $\overline{U}^{h\alpha}$, $R - U^{j\beta}$, $h, j \leqq i$, do not meet their distance $> 0$. We thus have an upper bound for $\epsilon_{i+1}$ the mesh of $\Sigma^{i+1}$ and we choose $\{\epsilon_1\} \longrightarrow 0$ monotonely. As a consequence of the construction, plus the last condition we verify that $\Delta$ thus obtained is normal.

3. The following additional conditions are readily imposed upon $\Delta$: (*a*) the order[1] of $\Sigma^{i+1}$ is the lowest possible for any $\Sigma \subset \Sigma^i$; in particular, if the order of $R$, say $\omega$, is finite we can take order $\Sigma^i = \omega$ for every $i$; (*b*) $F(U)$ is of dimension $< \eta = \dim R$. As a consequence and with unimportant modifications we can extend to $R$ the proof given by Urysohn for a compact space of

*The Lebesgue order theorem: dim $R$ = order $R$.*

Regarding this proposition, for the separable case it was only proved up to date that order $R \leqq \dim R$ (Menger).

4. Along with the construction of $\Delta$ we obtain an image $R$ on $H$ of our initial space having the following property: Let $R^*$ be the closure of $R$ on $H$. The set $R^*$ is self-compact and $\Delta$ induces on it a normal development $\Delta^* = \{\Sigma^{*i}\}$ such that $\overline{\Sigma}^{*i}$ is of the same order as $\overline{\Sigma}^i$ ($\overline{\Sigma}$ designates the aggregate of the closures of the sets of $\Sigma$). Let in particular the $\Sigma$'s be of order $\omega = n = \dim R$. Then likewise order $\Sigma^{*i} = n$ for every $i$. From this follows that $R^*$ is of order $\leqq n$ and hence that dim $R^* \leqq n$. But $R^*$ contains an $n$-set, namely, $R$ itself. Hence dim $R^* = n$ and we have

*Imbedding theorem.—R can always be mapped topologically on a compact metric space of the same dimension.*

In a recent paper[2] we have given a number of results regarding the immersion of a compact metric $n$-space in $S_{2n+1}$. *As a consequence of the imbedding theorem they are all valid for any separable metric space.*

[1] The *order* of an aggregate of sets $\{A\}$ is the maximum number $k$ such that there are groups of $k + 1$ intersecting $A$'s. Let $\Sigma$ be a finite covering by open sets and $\sigma$ another $\subset\Sigma$ whose order $k$ is the least possible. The upper bound $\omega$ of $k$ when $\Sigma$ varies is the *order* of $R$. The same number is obtained when the sets of $\sigma$ are closed or merely closures of open sets.

[2] *Ann. Math.*, **2**, 32, 521–538 (1931). Complete references to the other results mentioned in the present paper will be found in the recent article by Tietze and Vietoris in the *Encykl. der Math. Wiss.* (III, AB, 13).

# COMPLETE TRANSITIVITY AND THE ERGODIC PRINCIPLE

BY EBERHARD HOPF[1]

HARVARD COLLEGE OBSERVATORY

Communicated January 13, 1932

1. *Introduction.*—Let $P$ denote the points of a manifold $\Omega$, and let $m(A)$ be a measure in the sense of Lebesgue, defined for certain (measurable) point sets $A \subset \Omega$. Furthermore, let $T_t(P)$, $-\infty < t < +\infty$, be a one parameter group of one to one transformations of $\Omega$ into itself,

$$T_s T_t \equiv T_{s+t}, \quad T_0(P) \equiv P,$$

which preserve the measure $m$,

$$m(T_t(A)) = m(A).$$

It is supposed that

$$\lim_{t=0} m(A\,T_t(B)) = m(AB)$$

holds for any two measurable sets $A$, $B$. This continuity supposition is certainly fulfilled, if $\Omega$ is an analytic manifold and if $T_t(P)$ depends analytically upon $P$ and $t$.

We set

$$f_t(P) = f(T_t(P)),$$

$f(P)$ being a function of the points of $\Omega$. In this note we confine ourselves to q.s. (quadratically summable) functions of $P$. The usual notation

$$(f, g) = \int_\Omega f(P)\overline{g(P)}dm, \quad \|f\| = \sqrt{(f, f)}$$

is employed. B. O. Koopman[2] found that the transition from $f(P)$ to $f_t(P)$ represents for every $t$ a unitary transformation of the elements $f$ of the Hilbert space. This connection with the theory of the Hilbert space yields the spectral representation of the group $T_t(P)$,

$$(f_t, g) = \int_{-\infty}^{+\infty} e^{i\lambda t} d(E_\lambda f, g), \tag{1}$$

where the projection operators $E_\lambda$ satisfy the relations

$$E_{-\infty} = 0, E_{+\infty} f = f, \quad E_\lambda E_\mu = E_\mu \ (\lambda \geqq \mu), \quad E_{\lambda+0} = E_\lambda.$$

The time average

$$f^T(P) = \frac{1}{T}\int_0^T f_t(P)dt$$

exists almost everywhere on $\Omega$, $f(P)$ being a q. s. function. Recently v. Neumann[3] proved by means of (1) that any q. s. function $f(P)$ possesses a general time average $f^*(P)$ in the sense of mean convergence,

$$\lim_{T=\infty} ||f^T - f^*|| = 0.$$

An elementary proof has been given by the author.[4] By using entirely different methods Birkhoff[5] proved the remarkable fact that actually

$$\lim_{T=\infty} f^T(P) = f^*(P),$$

except in a set of points $P$ of measure zero.

The group $T_t(P)$ is called "metrically transitive," if $m(\Omega)$ is finite and if

$$m(A)m(\Omega - A) = 0$$

holds for any measurable point set which is invariant under all transformations $T_t(P)$ of the group,[6]

$$T_t(A) = A.$$

The notion of metric transitivity plays a fundamental rôle in the mathematical foundation of classical statistical mechanics. In the case of metric transitivity we have almost everywhere

$$f^*(P) = \text{const.} = \frac{\int_\Omega f(P)dm}{m(\Omega)}, \tag{2}$$

i.e., the time average is equal to the space average.[3]

2. *A Problem of Poincaré. Complete Transitivity.*—This paper is devoted to the solution of another ergodic problem (mèlange des liquides) considered by Poincaré.[7] In the case of three dimensions a group $T_t(P)$ with an invariant measure $m$ may be realized by a steady flow of an incompressible fluid contained in a vessel. Suppose that initially a certain part of the fluid is red colored. Experience shows then that "in general" after long time the red color is nearly uniformly distributed in the vessel.

The mathematical meaning of this fact is that, if the group $T_t(P)$ satisfies certain conditions (which will be stated below),

$$\lim_{|t|=\infty} (f_t, g) = (f^*, g) = \frac{\int_\Omega f dm \int g dm}{m(\Omega)}, \tag{3}$$

$f$, $g$ being any two q. s. functions, in other words, that for $t \longrightarrow \infty$ the (likewise for $t \longrightarrow -\infty$) functions $f_t(P)$ converge weakly (schwache Konvergenz). If, for instance, $f(P)$ represents the initial density distribution of the red-colored liquid, and if

$$g(P) = 1, \quad P \subset A; \quad g(P) = 0, \quad P \subset \Omega - A,$$

then $(f_t, g)$ indicates the exact amount of red-colored liquid which is contained in $A$ at the time $-t$, and we have $(f^*, g) = \text{const. } m(A)$, i.e., we have uniform mixture for $|t| \longrightarrow \infty$.

We remark first that the condition of metric transitivity is *not* sufficient in establishing (3). Let $\varphi$, $\vartheta$ be the angular coördinates of the points $P$ of the torus. Poincaré[7] recognized that the group of the transformations

$$T_t(P) : \varphi, \vartheta \longrightarrow \varphi + \alpha t, \vartheta + \beta t; \quad \alpha/\beta \text{ irrational}, \tag{4}$$

of the torus into itself has not the property (3) though being metrically transitive.[4]

The notion of metric transitivity may be associated with a single transformation $T$ of $\Omega$ into itself preserving the measure $m$. $T$ is metrically transitive, if

$$m(A)m(\Omega - A) = 0$$

follows from $A = T(A)$, $A$ being a measurable set. Now a necessary condition for (3) is that any transformation $T_t$ of the group is metrically transitive (except, of course, the identity $T_0$). Such a group may be called "completely transitive." Groups $T_t$ probably have this property in general. Complete transitivity is certainly a consequence of (3). Indeed,

$$T_a(A) = A, \quad T_t(A) \not\equiv A, \quad 0 < m(A) < m(\Omega)$$

implies $T_{t+a}(A) = T_t(A)$. Hence, in setting

$$f(P) = 1, \quad P \subset A; \; f(P) = 0, \quad P \subset \Omega - A,$$

we infer that, for a suitable $g(P)$,

$$(f_t, g) = \int_B g(P)dm, \quad B = T_{-t}(A),$$

is a periodic function of $t$ without being constant.

I was not able to prove that complete transitivity is sufficient for (3), although I have little doubt about its being true. As a first approximation to the truth, however, the following theorem may be communicated:

*Theorem* 1.—Complete transitivity is a necessary and sufficient condition that

$$\lim_{|T|=\infty} \frac{1}{T} \int_0^T \left| (f_t, g) - \frac{\int_\Omega f dm \int_\Omega g dm}{m(\Omega)} \right| dt = 0$$

holds for any two q. s. functions $f$, $g$.

It remains to prove that the condition is sufficient. The function $(E_\lambda f, g)$ is known to be of bounded variation over the whole of $-\infty < \lambda < +$

$\infty$. Let $\lambda_0 = 0$, $\lambda_1$, $\lambda_2$, ... be the characteristic values forming the pure point spectrum of (1). We have

$$(f_t, g) = \phi_1(t) + \phi_2(t),$$
$$\phi_1 = \sum_r e^{i\lambda_r t}(\varphi_r, g), \quad \varphi_r = (E_{\lambda_r} - E_{\lambda_r - 0})f, \tag{5}$$

where $\phi_2$ is the Fourier transform of a continuous function of bounded variation. Such a function is readily seen (see section 3) to have the property

$$\lim_{T=\infty} \frac{1}{T} \int_0^T |\phi_2(t)|\, dt = 0. \tag{6}$$

The $\varphi_r(P)$ are characteristic functions,

$$\varphi_r(T_t(P)) = e^{i\lambda_r t}\varphi_r(P), \quad -\infty < t < +\infty.$$

If we call the sum of an almost periodic function and of a function with the property (6) an "asymptotically almost periodic function," we infer that $(f_t, g)$ is a function of this kind.

It is an important question whether the continuous part of the function $(E_2 f, g)$ is absolutely continuous if the spectrum is associated with a steady flow $T_t(P)$. If this is true we would have $\phi_2(t) \longrightarrow 0$ for $|t| \longrightarrow \infty$ and $(f_t, g)$ would be the sum of an almost periodic function and a function that tends to zero if $|t| \longrightarrow \infty$, in particular, $m(AT_t(B))$ would have this property, $A$ and $B$ being any two sets.[8]

According to v. Neumann, $\varphi_0(P)$ is the time average $f^*(P)$ of $f(P)$. As complete transitivity implies metric transitivity of the group $T(P)$, (2) gives the value of the time average. Furthermore, any measurable function which is invariant under a transformation $T_a$, $a \neq 0$, must be constant almost everywhere on $\Omega$. Hence we conclude that the point spectrum contains only one point, $\lambda = 0$. Indeed, $\varphi_t(P) = e^{i\lambda t}\varphi(P)$, $\varphi \not\equiv 0$, implies

$$\varphi(T_a(P)) = \varphi_a(P) = \varphi(P), \quad a = \frac{2\pi}{\lambda}, \quad \varphi(P) \not\equiv \text{const.}$$

We therefore have

$$\phi_1(t) = \text{const.} = (\varphi_0, g) = \frac{\int_\Omega f dm \int_\Omega g dm}{m(\Omega)}, \tag{7}$$

whence the theorem follows immediately.

The following theorem may be mentioned in connection with the spectral theory.

*Theorem.*—Complete transitivity of the group $T_t(P)$ is the necessary and sufficient condition that $\lambda = 0$ is a simple and the only characteristic value of the spectrum of the group.

The proof is implicitly included in the above considerations. A characteristic function belonging to a characteristic value $\lambda \neq 0$ cannot exist in the case of complete transitivity. Conversely, if $\lambda = 0$ is a simple and the only characteristic value, the equations (5), (6) and (7) hold, and the complete transitivity follows from theorem 1.

3. *Proof of* (6).—It is sufficient to prove that a function of the form

$$\phi(t) = \int_{-\infty}^{+\infty} e^{i\lambda t} d\varphi(\lambda),$$

where $\varphi(\lambda)$ is continuous, non-decreasing and bounded in $(-\infty, +\infty)$, has the property

$$\lim_{T=\infty} \frac{1}{T} \int_0^T |\phi(t)| dt = 0. \tag{7}$$

We have

$$\frac{1}{2}|\phi(t)|^2 = \int\int \cos(\lambda - \mu) t d\varphi(\lambda) d\varphi(\mu),$$

thus yielding

$$\frac{1}{2T} \int_0^T |\phi|^2 dt = \int\int \frac{\sin(\lambda - \mu)T}{(\lambda - \mu)T} d\varphi(\lambda) d\varphi(\mu).$$

This integral may be split up into an integral $I_1$ over the strip $|\lambda - \mu| < \delta$ and the remaining integral $I_2$. Obviously we have

$$|I_2| \leqq \frac{1}{\delta T} [\varphi(+\infty) - \varphi(-\infty)]^2.$$

On the other hand, from $|\sin x| \leqq |x|$,

$$|I_1| \leqq \int_{-\infty}^{+\infty} [\varphi(\lambda + \delta) - \varphi(\lambda - \delta)] d\varphi(\lambda)$$

$$\leqq \max_{\lambda} [\varphi(\lambda + \delta) - \varphi(\lambda - \delta)] \cdot [\varphi(+\infty) - \varphi(-\infty)],$$

thus yielding $I_1 + I_2 \longrightarrow 0$ if $T \longrightarrow \infty$. Schwartz's inequality

$$\left(\frac{1}{T} \int_0^T |\phi| dt\right)^2 \leqq \frac{1}{T} \int_0^T |\phi|^2 dt$$

completes the proof of (6).

Hitherto, no group $T_t$ has been proved to be completely transitive. The author expects to treat certain examples suspected already by Poincaré of having the property (3).

[1] International Research Fellow.

[2] B. O. Koopman, these PROCEEDINGS, **17**, 315–318 (1931).

[3] J. von Neumann, these PROCEEDINGS, **18,** 70–82 (1932).
[4] E. Hopf, these PROCEEDINGS, **18,** 93–100 (1932).
[5] G. D. Birkhoff, these PROCEEDINGS, **17,** 656–660 (1931).
[6] G. D. Birkhoff and P. Smith, *Journal de Math.*, **7** [9],(1928), p. 365.
[7] H. Poincaré, *Calcul des Probabilitiés*, 2nd Edition, Paris, 1912, 320–333.
[8] This remark is due to B. O. Koopman who privately communicated it to the author two months ago.

---

## *THE INCUBATING PYTHON: A TEMPERATURE STUDY*

BY FRANCIS G. BENEDICT, EDWARD L. FOX AND V. COROPATCHINSKY

NUTRITION LABORATORY OF THE CARNEGIE INSTITUTION OF WASHINGTON, BOSTON, MASSACHUSETTS

Read before the Academy, November 16, 1931

It is the fixed belief of every farmer's wife that the sitting and the broody hen have high body temperatures. That they are fussy and agitated no one doubts, but in 1911 Simpson[1] showed that sitting and broody birds do not have high temperatures. Singularly enough, this erroneous conception of a feverish state in the sitting or brooding bird has persisted for decades and has led to the conviction that the python coils about her eggs in order to supply heat to incubate them. The authentic history of the incubating python begins with a report from the French naturalist, Lamarre-Picquot[2] who in 1832 stated before the French Academy that the python of India incubates her eggs with sensible heat from the body. Since this statement was accompanied by a story to the effect that there existed likewise a snake that could "withdraw milk from the udder of a cow," his whole communication was given little credence. Indeed, the sharpest critics in the French Academy rejected the statements of Lamarre-Picquot as being hazardous and questionable. There then followed a controversy on the subject, which lasted many years. In 1841 Valenciennes[3] presented some experimental evidence obtained in studying a python that had deposited her eggs in the Jardin des Plantes in Paris and had coiled about them, as usual. Although the method of controlling the temperature in a snake cage was crude at that time in comparison with the more modern methods, Valenciennes tried by every means to rule out all other factors but that of the body heat of the snake, and he was convinced that the python had a high temperature. In this belief he was strongly opposed by the Director of the Jardin des Plantes at that time, Duméril,[4] who contended that in the particular case cited by Valenciennes the eggs were spoiled and the heat produced was in large part due to the fermentation of the eggs themselves. Subsequent observations made in the London Zoölogical Park, first by

Sclater[5] in 1862 and later by Forbes[6] in 1881, seemingly confirmed Valenciennes' observations, but the difficulty of comparing the temperature of the animal and the temperature of the air was always uppermost. Both Sclater and Forbes measured the body temperature not only of a male but of a female python and found that at times the male's temperature was higher than that of the female. In general, however, the female's temperature was above that of the male, and the temperatures of both animals were for the most part above the room temperature.

In a study of the metabolism and the mechanism of heat production of snakes made by the Nutrition Laboratory, chiefly at the New York Zoölogical Park but with some direct calorimetric measurements in Boston, we were led to believe that *under ordinary conditions* the snake has a body temperature slightly below its environmental temperature[7] and that all the heat produced by the snake is lost by vaporization of water. Under two conditions only was the temperature of the snake found to be consistently above that of the environment. At the peak of digestion the snake's temperature was slightly higher than the temperature of the air around it, and after pronounced agitation it was considerably increased. It became important therefore to determine whether a snake, unaided by the process of digestion or by agitation, can operate any mechanism for raising its temperature above the temperature of its environment. Quite by chance it was found that a python was incubating her eggs in the new Reptile House in the National Zoölogical Park at Washington, D. C., and through the kindness of the Director, Dr. W. M. Mann, the Nutrition Laboratory was permitted to make a study of this animal. An electrical equipment consisting of galvanometer, thermo-elements and thermostats was sent to the Park, and since it was realized that it is as difficult to determine accurately the temperature of the environment as that of the animal itself, an experimental procedure was planned that enabled the determination not only of the animal's temperature but of the environmental temperature.

Fortunately the python was most placid and would allow any amount of handling with hardly any indication of disturbance. In the entire series of experiments with this animal she gave but one sign of agitation, which was a heavy, deep breath, approximating a hiss, in the middle of an afternoon session. In any large den in a Zoölogical Park there is always a possible source of heat from water pipes and from sunlight entering through the windows and skylight, and there is possibility of a cooling effect from the contact of the glass window of the den with the cool air of the corridor. One could hardly expect, therefore, that the temperature of the air surrounding the snake in such a den would be equable. Hence, a temperature survey of the air about the python was first made, the animal being considered to be in a hypothetical, rectangular box of air, and the gradients in the temperature of this air at different distances above the floor and away

from the snake being carefully noted. Temperature measurements were then made at various points on the exposed surface of the snake, between the folds, between the snake and the gravel on the floor of the den, and occasionally between the folds and the eggs. Since the heating supply of the building was entirely shut off, there was no heat under the gravel.

Our preliminary measurements showed that the skin temperature of the snake was definitely above the environmental temperature. The importance of this observation decided us to make three entirely independent temperature surveys, after each of which the instruments were dismantled, the thermostats reset, and the galvanometer and the thermo-junctions recalibrated. An abstract of the results is given in table 1. The actual number of readings represented by each average temperature is recorded. The first half of the table contains the average air temperatures as measured on the graveled floor near the python, and 10, 30 and 60 cm. in a vertical line above the floor, these vertical gradients being determined at four different points about the snake. The second section of the table contains the skin temperature.

TABLE 1

COMPARISON OF ENVIRONMENTAL AND SKIN TEMPERATURES OF AN INCUBATING PYTHON

| MEASUREMENT | SURVEY 1 NO. OF READINGS | SURVEY 1 °C. | SURVEY 2 NO. OF READINGS | SURVEY 2 °C. | SURVEY 3 NO. OF READINGS | SURVEY 3 °C. | AVERAGE TEMPERATURE[1] |
|---|---|---|---|---|---|---|---|
| Air temperatures: | | | | | | | |
| Gravel near python | 27 | 31.38 | 3 | 32.49 | 8 | 31.08 | 31.65 |
| 10–15 cm. above floor | 20 | 31.01 | 8 | 30.23 | 9 | 30.67 | 30.64 |
| 30 cm. above floor | 19 | 31.07 | 8 | 30.39 | 5 | 29.41 | 30.29 |
| 60 cm. above floor | 9 | 31.24 | 3 | 30.44 | 4 | 29.28 | 30.32 |
| Grand average | | | | | | | *30.73* |
| Skin temperatures: | | | | | | | |
| Under python | 10 | 33.15 | 15 | 33.32 | 10 | 33.23 | 33.23 |
| Near gravel | 6 | 33.00 | .. | ... | 4 | 33.19 | 33.10 |
| On exposed surface | 33 | 33.73 | 16 | 34.31 | 17 | 33.90 | 33.98 |
| Between folds | 15 | 33.85 | 10 | 34.61 | 13 | 34.81 | 34.42 |
| Grand average | | | | | | | *33.68* |

[1] Not weighted according to number of readings taken.

There was a singular uniformity in the temperature of the air at the various parts of the hypothetical air box. In spite of the cool air of the corridor outside the glass window of the den, the high temperature of the air inside the den, and the presence of the observer (although his activity was kept at a minimum), the temperature of the air about the snake was reasonably constant, averaging for the three surveys 30.73°C. The skin temperature of the python was highest between the folds. These temperatures should perhaps more properly be designated as the "body

temperature," since there was no chance for vaporization between the folds and a higher temperature here is to be expected. The average of all the skin temperature measurements is 33.68°C. There was, therefore, a difference of almost exactly 3°C. between the average temperature of the air about the python and the temperature of the animal itself. If one considers that the rectal temperature (had it been taken) would be more nearly approximated by the temperature between the folds, which was found to average 34.42°C., the difference between the air and the body temperatures would be almost 4°C.

Under the head of "air temperatures" in table 1 we have included the temperatures of the gravel near the python. It is more than likely that these temperatures were influenced by the higher temperature of the animal itself and there were convection currents, although rather feeble. If the temperatures of the gravel are not taken into consideration, the average environmental temperature becomes somewhat lower, averaging 30.4°C. This temperature is a full 4° below the temperature between the snake's folds, that is, the probable body temperature. Although subsequently the python's eggs were found to be infertile and partially decomposed, it is extremely improbable that such bacterial action can account for more than a very small percentage of the sensible heat noted in this study.

This observation, that the incubating python has a body temperature higher than the air surrounding it, bids fair to be of considerable importance in comparative physiology. Since in the various stages of animal development from the cold-blooded or poikilotherm to the warm-blooded or homoiotherm the birds represent the next advancement after the reptiles, and since birds invariably incubate their eggs, it is possible that the incubating python forms an intermediary stage between the non-incubating reptile and the bird.

The details of this study are soon to appear in a monograph on "The Physiology of Large Reptiles" (Carnegie Institution of Washington Publication No. 425), which is now in press.

[1] Simpson, S., *Proc. Roy. Soc. Edinburgh*, **32** (Pt. 1), No. 5, p. 25, and No. 11, p. 131, 1911–1912.

[2] Lamarre-Picquot, *L'Institut*, **3**, p. 70, 1835.

[3] Valenciennes, A., *Compt. Rend.*, **13**, p. 126, 1841.

[4] Duméril, *Ibid.*, **14**, p. 193, 1842.

[5] Sclater, P. L., *Proc. Zoöl. Soc., London*, p. 365, 1862.

[6] Forbes, W. A., *Ibid.*, p. 960, 1881.,

[7] Benedict, F. G., and E. L. Fox, *Proc. Nat. Acad. Sci.*, **17**, p. 584, 1931.